

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 19, 2004, 10:01:13 ; Search time 41 Seconds

(without alignments)
42.585 Million cell updates/sec

Title: US-09-551-336B-1

Perfect score: 45

Sequence: 1 AAAAAAAAAAK 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	16	AAW32224	Beta-sheet forming
2	45	100.0	16	AAW32226	Beta-sheet forming
3	45	100.0	17	AAW32227	Beta-sheet forming
4	45	100.0	20	ABP59409	Self-assembling di
5	45	100.0	30	ABP59422	Self-assembling tr
6	45	100.0	30	ABP59423	Self-assembling tr
7	45	100.0	97	ABG75920	Human colon cancer
8	45	100.0	201	ABP41465	Human ovarian anti
9	45	100.0	216	ABP62982	Human polypeptide

10	45	100.0	218	21	AAW00755
11	45	100.0	218	21	AAW00759
12	45	100.0	220	23	AAU76972
13	45	100.0	220	24	ABU07443
14	45	100.0	221	23	ABP62855
15	45	100.0	225	21	AAW00758
16	45	100.0	226	17	AAW05151
17	45	100.0	262	22	ABG15586
18	45	100.0	265	21	ABP58221
19	45	100.0	335	22	ABP58955
20	45	100.0	406	22	ABP58251
21	45	100.0	515	21	AAW6135
22	45	100.0	571	21	AAV69071
23	45	100.0	660	20	AAV01303
24	45	100.0	698	20	AAV01302
25	45	100.0	698	21	AAV69069
26	45	100.0	712	21	AAW08630
27	45	100.0	730	19	AAW46315
28	45	100.0	730	21	AAW08631
29	45	100.0	730	23	AAO17360
30	45	100.0	731	21	AAV69068
31	45	100.0	731	22	ABP6657
32	45	100.0	733	15	AAW5653
33	45	100.0	733	20	AAV01301
34	42	93.3	20	24	ABP59411
35	42	93.3	30	24	ABP59424
36	42	93.3	30	24	ABP59425
37	42	93.3	31	21	AAW08166
38	42	93.3	102	24	ABU20243
39	42	93.3	102	24	ABU20253
40	42	93.3	102	24	ABU20256
41	42	93.3	157	22	AAU16372
42	42	93.3	157	24	ABU5441
43	42	93.3	161	21	AAW19319
44	42	93.3	229	23	ABP41248
45	42	93.3	244	22	ABG21054

ALIGNMENTS

Human secreted pro
Human secreted pro
Human ribosomal L1
Protein different
Human polypeptide
Human secreted pro
Nucleic proliferat
Novel human diagno
Lung cancer associ
Drosophila melanog
Human protein sequ
Amino acid sequenc
Human tropoelastin
Human tropoelastin
Amino acid sequenc
Amino acid sequenc
Human elastin prot
Human elastin prot
Fusion protein com
Human elastin. Ho
Amino acid sequenc
Human elastin
Synthetic human tr
Amino acid sequenc
Self-assembling di
Self-assembling tr
Peptide modulating
162P186 cancer gen
162P186 cancer gen
Human novel secret
Human novel polype
Amino acid sequenc
Human ovarian anti
Novel human diagno

RESULT 1	AAW32224	ID	AAW32224 strand; peptide; 16 AA.
AC	AAW32224;	XX	
DT	12-FEB-1998 (first entry)	XX	
DE	Beta-sheet forming peptide #1.	XX	
KW	Beta sheet; circular dichroism spectroscopy; glycosidase activity;	XX	
KW	phosphodiesterase activity; drug screening; hydrolysis;	XX	
KW	neurodegenerative disease; Alzheimer's disease; amyloid deposition.	XX	
OS	Synthetic.	XX	
FT	Key	FT	Location/Qualifiers
FT	Modified-site	FT	1 /note= "N-terminal acetyl"
FT	Modified-site	FT	16 /note= "C-terminal amide"
PN	W09637212.A1.	PN	
XX	28-NOV-1996.	XX	
PD	23-MAY-1996;	PD	96WO-US07564.
PF	26-MAY-1995;	PF	95US-0452043.
PR	(TORR-) TORREY PINES INST MOLECULAR STUDIES.	PR	
XX		XX	
PA		PA	

PI Blondelle SE, Forood B, Houghten RA, Perez-paya E;
XX
XX MPI; 1997-033944/03.

PT Polypeptide(s) which form stable beta sheets in aq. environment -
PT having phosphodiesterase and glycosidase activity and useful in
PT bio:engineering, enzymatic and drug screening applications

PS Example 1; Page 14; 35pp; English.

XX This sequence represents a specific example of a polypeptide, which
CC forms a beta sheet in an aqueous environment and has the general formula
CC JNamXum20n or Ac-JNamXum20n-NH2, in which A = D- or L-alanine
CC amino acid; m1, m2 = 0-40; m1+m2 = 10-40; J = a charged amino acid; n = 1
CC or 2; X = any amino acid except proline; u = 0 or 1; yl lysine,
CC hydroxylysine, arginine, histidine, aspartic acid, glutamic acid or
CC gamma-carboxyglutamic acid; Ac = an acylation modification to the amino
CC terminus; NH2 = an amidation modification to the carboxylic acid
CC terminus; and X1 = cysteine, threonine, tyrosine or serine. The peptides
CC have phosphodiesterase activity and readily hydrolyse nucleic acids.
CC They also have glycosidase activity and readily hydrolyse sugar
CC phosphates. Further they can cause amine-catalysed decarboxylation (e.g.
CC decarboxylation of oxalacetate), and they can hydrolyse phospholipids.
CC They can be used for bioengineering, enzymatic and drug screening
CC applications. In particular they can be used for screening drugs for the
CC prevention or treatment of neurodegenerative diseases, e.g. Alzheimer's
CC disease, which result in amyloid protein deposition. They can also be
CC used to screen for drugs which inhibit or disrupt the beta sheet.

XX Sequence 16 AA;

Query Match 100.0%; Score 45; DB 18; Length 16;

Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
| | | | | | | | | |
Db 6 AAAAAAAAAAK 16

RESULT 2
AAW32226

ID AAW32226 standard; peptide; 16 AA.

XX AAW32226;

DT 12-FEB-1998 (first entry)

XX Beta-sheet forming peptide #2.

XX Beta sheet; circular dichroism spectroscopy; glycosidase activity;
KM phosphodiesterase activity; drug screening; hydrolysis;
KW neurodegenerative disease; Alzheimer's disease; amyloid deposition.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 16 /note= "C-terminal amide"

XX W0963 212-A1.

XX 28-NOV-1996.

XX 23-MAY-1996; 96WO-US07564.

XX 26-MAY-1995; 95US-0452043.

XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.

XX Blondelle SE, Forood B, Houghten RA, Perez-paya E;

XX MPI; 1997-033944/03.

PT Polypeptide(s) which form stable beta sheets in aq. environment -
PT having phosphodiesterase and glycosidase activity and useful in
PT bio:engineering, enzymatic and drug screening applications

PS Example 1; Page 15; 35pp; English.

XX This sequence represents a specific example of a polypeptide, which
CC forms a beta sheet in an aqueous environment and has the general formula
CC JNamXum20n or Ac-JNamXum20n-NH2, in which A = D- or L-alanine
CC amino acid; m1, m2 = 0-40; m1+m2 = 10-40; J = a charged amino acid; n = 1
CC or 2; X = any amino acid except proline; u = 0 or 1; yl lysine,
CC hydroxylysine, arginine, histidine, aspartic acid, glutamic acid or
CC gamma-carboxyglutamic acid; Ac = an acylation modification to the amino
CC terminus; NH2 = an amidation modification to the carboxylic acid
CC terminus; and X1 = cysteine, threonine, tyrosine or serine. The peptides
CC have phosphodiesterase activity and readily hydrolyse nucleic acids.
CC They also have glycosidase activity and readily hydrolyse sugar
CC phosphates. Further they can cause amine-catalysed decarboxylation (e.g.
CC decarboxylation of oxalacetate), and they can hydrolyse phospholipids.
CC They can be used for bioengineering, enzymatic and drug screening
CC applications. In particular they can be used for screening drugs for the
CC prevention or treatment of neurodegenerative diseases, e.g. Alzheimer's
CC disease, which result in amyloid protein deposition. They can also be
CC used to screen for drugs which inhibit or disrupt the beta sheet.

XX Sequence 16 AA;

Query Match 100.0%; Score 45; DB 18; Length 16;

Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
| | | | | | | | | |
Db 6 AAAAAAAAAAK 16

RESULT 3
AAW32227

ID AAW32227 standard; peptide; 17 AA.

XX AAW32227;

DT 12-FEB-1998 (first entry)

XX Beta-sheet forming peptide #3.

XX Beta sheet; circular dichroism spectroscopy; glycosidase activity;
KM phosphodiesterase activity; drug screening; hydrolysis;
KW neurodegenerative disease; Alzheimer's disease; amyloid deposition.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 17 /note= "C-terminal amide"

XX W0963 212-A1.

XX 28-NOV-1996.

XX 23-MAY-1996; 96WO-US07564.

XX 26-MAY-1995; 95US-0452043.

XX (TORR-) TORREY PINES INST MOLECULAR STUDIES.

XX Blondelle SE, Forood B, Houghten RA, Perez-paya E;

WPI; 1997-033944/03.

Polypeptide(s) which form stable beta sheets in aq. environment -
bioengineering, enzymatic and drug screening applications

Example 2; Page 16; 35pp; English.

This sequence represents a specific example of a polypeptide, which forms a beta sheet in an aqueous environment and has the general formula JnAmXkAm2n or $\text{Ac-JnAmXkAm2n-NH}_2$, in which A = D- or L-alanine or amino acid; m1, m2 = 0-40; m1+m2 = 10-40; J = a charged amino acid; n = 1 or 2; X = any amino acid except proline; u = 0 or 1; u1 lysine, gamma-oxyllysine, arginine, histidine, aspartic acid, glutamic acid or gamma-carboxyglutamic acid; Ac = an acylation modification to the amino terminus; NH2 = an amidation modification to the carboxylic acid terminus; and X1 = cysteine, threonine, tyrosine or serine. The peptides have phosphodiesterase activity and readily hydrolyse nucleic acids. They also have glycosidase activity and readily hydrolyse sugar phosphates. Further they can cause amine-catalysed decarboxylation (e.g. decarboxylation of oxalacetate), and they can hydrolyse phospholipids. They can be used for bioengineering, enzymatic and drug screening applications. In particular they can be used for screening drugs for the prevention or treatment of neurodegenerative diseases, e.g. Alzheimer's disease, which result in amyloid protein deposition. They can also be used to screen for drugs which inhibit or disrupt the beta sheet.

Sequence 17 AA;

Query Match 100.0%; Score 45; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
|||||||
6 AAAAAAAAAAK 16

RESULT 4
ABP59409

ID ABP59409 standard; peptide; 20 AA.

ABP59409;

09-JUN-2003 (first entry)

Self-assembling di-block oligopeptide AK20.

Self-assembling oligopeptide; dipolar; nanotube; nanostructure;
drug delivery; carrier.

Synthetic.

WO2003006043-A1.

23-JAN-2003.

10-JUL-2002; 2002WO-US21757.

10-JUL-2001; 2001US-304256P.

(MASI) MASSACHUSETTS INST TECHNOLOGY.

Zhang S, Vauchey S;

WPI; 2003-267944/26.

New dipolar oligopeptides in a self-assembled nanostructure useful for
drug delivery

Claim 16; Page 21; 52pp; English.

The present invention relates to self-assembling dipolar oligopeptides

and di- and tri-block peptide copolymers. The dipolar oligopeptides have ability to self assemble to form stable nanotubes. The self-assembled nanostructure have the ability to entrap and deliver molecules with high degree of efficacy. The copolymers are amenable for molecular systematic design, modification and simulations before synthesis; can be highly purified to be mono-dispersed materials; combinatorial approach can be employed to systematically characterize these co-polymers at various ratio; and can be synthesized in vitro or in vivo. The nanotubes fuse with the lipid bilayers and do not deform the cells as liposomes do. The oligopeptides can be used in a self-assembled nanostructure (e.g. nanotube) for delivering drug into a cell and as a carrier for biologically active materials. The present sequence is one such di-block oligopeptide.

Sequence 20 AA;

Query Match 100.0%; Score 45; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
|||||||
1 AAAAAAAAAAK 11

RESULT 5
ABP59422

ID ABP59422 standard; peptide; 30 AA.

ABP59422;

09-JUN-2003 (first entry)

Self-assembling tri-block oligopeptide KAK30.

Self-assembling oligopeptide; dipolar; nanotube; nanostructure;
drug delivery; carrier.

Synthetic.

WO2003006043-A1.

23-JAN-2003.

10-JUL-2002; 2002WO-US21757.

10-JUL-2001; 2001US-304256P.

(MASI) MASSACHUSETTS INST TECHNOLOGY.

Zhang S, Vauchey S;

WPI; 2003-267944/26.

New dipolar oligopeptides in a self-assembled nanostructure useful for
drug delivery

Claim 21; Page 23; 52pp; English.

The present invention relates to self-assembling dipolar oligopeptides and di- and tri-block peptide copolymers. The dipolar oligopeptides have ability to self assemble to form stable nanotubes. The self-assembled nanostructure have the ability to entrap and deliver molecules with high degree of efficacy. The copolymers are amenable for molecular systematic design, modification and simulations before synthesis; can be highly purified to be mono-dispersed materials; combinatorial approach can be employed to systematically characterize these co-polymers at various ratio; and can be synthesized in vitro or in vivo. The nanotubes fuse with the lipid bilayers and do not deform the cells as liposomes do. The oligopeptides can be used in a self-assembled nanostructure (e.g. nanotube) for delivering drug into a cell and as a carrier for

CC biologically active materials. The present sequence is one such
CC tri-block oligopeptide.

XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 45; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.56;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11

Db 11 AAAAAAAAAAK 21

RESULT 6

ABP59423 ID ABP59423 standard; peptide; 30 AA.

XX AC ABP59423;

XX DT 09-JUN-2003 (first entry)

XX DE Self-assembling tri-block oligopeptide AKA30.

XX KM Self-assembling oligopeptide; dipolar; nanotube; nanostructure;
XX drug delivery; carrier.

XX OS Synthetic.

XX PN WO200306043-A1.

XX PD 23-JAN-2003.

XX PF 10-JUL-2002; 2002WO-US21757.

XX PR 10-JUL-2001; 2001US-304256P.

XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX PI Zhang S, Vauthey S;

XX DR WPI; 2003-267944/26.

XX PT New dipolar oligopeptides in a self-assembled nanostructure useful for
XX drug delivery -

XX PS Claim 23; Page 23; 52pp; English.

XX CC The present invention relates to self-assembling dipolar oligopeptides
XX and di-and tri-block peptide copolymers. The dipolar oligopeptides have
XX ability to self assemble to form stable nanotubes. The self-assembled
XX nanostructure have the ability to entrap and deliver molecules with high
XX degree of efficacy. The copolymers are amenable for molecular systematic
XX design, modification and synthesis; can be subjected to extensive
XX molecular modelling and simulations before synthesis; can be highly
XX purified to be mono-dispersed materials; combinatorial approach can be
XX employed to systematically characterize these co-polymers at various
XX ratios; and can be synthesized in vitro or in vivo. The nanotubes fuse
XX with the lipid bilayers and do not deform the cells as liposomes do.
XX The oligopeptides can be used in a self-assembled nanostructure (e.g.
XX nanotube) for delivering drug into a cell and as a carrier for
XX biologically active materials. The present sequence is one such
XX tri-block oligopeptide.

XX SQ Sequence 30 AA;

Query Match 100.0%; Score 45; DB 24; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11

Db 1 AAAAAAAAAAK 11

RESULT 7

AAAG75920 ID AAAG75920 standard; Protein; 97 AA.

XX AC AAAG75920;

XX DT 03-SEP-2001 (first entry)

XX DE Human colon cancer antigen protein SEQ ID NO:6684.

XX KM Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US26524.

XX PR 29-SEP-1999; 99US-0157137.

XX PR 03-NOV-1999; 99US-0163280.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX DR WPI; 2001-235357/24.

XX DR N-PSDB; AAH35325.

XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -

XX PS Claim 11; Page 8150; 9803pp; English.

XX CC AAH32943 to AAH37195 and AAAG73514 to AAAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing
XX inactive proteins or to supplement the patient's own production of P.
XX Additionally, N may be used to produce the colon cancer-associated P,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAAG7789 represent sequences used in the exemplification of the
XX present invention.

XX CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 97 AA;

Query Match 100.0%; Score 45; DB 22; Length 97;

Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11

Db 77 AAAAAAAAAAK 87

RESULT 8

ABPA1465 ID ABPA1465 standard; Protein; 201 AA.

XX

AC ABP41465;
XX 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HOO0038, SEQ ID NO:2597.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX
XX Homo sapiens.
OS
XX WO200200677-A1.
XX
XX 03-JAN-2002.
XX
XX 07-JUN-2001; 2001WO-US18569.
XX
XX 07-JUN-2000; 2000US-209467P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
PI
XX WPI: 2002-147878/19.
XX
XX N-PSDB; ABQ54542.
DR
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
XX
PS Claim 11; SEQ ID NO 2597; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 201 AA;
SQ

Query Match 100.0%; Score 45; DB 23; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAAK 11
DB 136 AAAAAAAAAAK 146
RESULT 9
ABP62982
ID ABP62982 standard; Protein, 216 AA.
AC ABP62982;
XX
XX 14-OCT-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 419.
DE
XX
XX Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
XX Homo sapiens.
OS
XX WO200218424-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US27093.
XX
XX 01-SEP-2000; 2000US-0654935.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehman T;
XX
XX WPI: 2002-583321/62.
XX
XX N-PSDB; ABQ93461.
DR
XX
XX New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies -
XX
XX
PS Claim 20; SEQ ID NO 419; 284pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising one of
XX 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
XX administering to a mammalian subject a composition comprising the protein
XX (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
XX (I), (II) and (III) are useful for diagnostic evaluation of disorders.
XX (I) is useful for gene therapy of diseases and (II) can be used for
XX therapeutic treatment. Diseases that may be treated include wound healing
XX and tissue repair, burns, central nervous system disorders (e.g.
XX Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
XX sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
XX sclerosis, diabetes and allergies.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 216 AA;
SQ
Query Match 100.0%; Score 45; DB 23; Length 216;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAAK 11

Db 151 AAAAAAAAAAK 161

RESULT 10

AAAG00755 ID AAG00755 standard; Protein; 218 AA.

AC AAG00755;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 4836.

KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KM gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PS (GSEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC00761.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 4836; 71bp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 218 AA;

Query Match 100.0%; Score 45; DB 21; Length 218;

Best Local Similarity 100.0%; Pred. NO. 3.6; 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11

Db 155 AAAAAAAAAAK 165

RESULT 11

AAAG00759 ID AAG00759 standard; Protein; 218 AA.

AC AAG00759;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 4840.

KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KM gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PS (GSEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC00765.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 4840; 71bp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 218 AA;

Query Match 100.0%; Score 45; DB 21; Length 218;

Best Local Similarity 100.0%; Pred. NO. 3.6; 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11

Db 155 AAAAAAAAAAK 165

RESULT 12

AAU76972 ID AAU76972 standard; Protein; 220 AA.

AC AAU76972;

DT 21-MAY-2002 (first entry)

DE Human ribosomal L14 (RPL14) protein.

KM Ribosomal L14 protein; RPL14; CD39L3; PMGM; GC20; cancer; metastasis;

KM carcinoma; non-small cell carcinoma; smoking; lung cancer;

KM bladder cancer; head cancer; neck cancer; urothelial cancer;

KM kidney cancer; pancreas cancer; mouth cancer; throat cancer;

KM pharynx cancer; larynx cancer; upper airway primary cancer;

KM upper airway secondary cancer; esophagus cancer; chromosome 3p21.3.

OS Homo sapiens.

PN WO200212563-A2.

14-FEB-2002.
14-AUG-2001; 2001WO-US24718.
06-AUG-2001; 2000US-222811P.
04-AUG-2000; 2000US-222811P.
(TEXA) UNIV TEXAS SYSTEM.
Katz R, Jiang F;
WPI; 2002-217200/27.
N-PSDB; ABK10349.
Identifying subject at risk for development of cancer, preferably lung cancer, comprises contacting RPL14, CD39L3, PMGM, or GC20 gene probe with test sample obtained from subject, and analysing DNA from test sample -
Example 1, Page 69-70; 79pp; English.
The invention describes a method of identifying a subject at risk for the development of cancer, predicting progression or metastasis of non-small cell carcinoma and other carcinoma in a subject, or identifying an individual to be segregated from a high risk environment. The method comprises contacting an RPL14, CD39L3, PMGM, or GC20 gene probe with a test sample obtained from a subject, and analysing DNA from the test sample. The method is useful for identifying a subject (a smoker, non-smoker or former smoker) at risk for the development, recurrence, or metastasis of cancer (preferably cancer of lung, bladder, head, neck, urothelial, kidney, pancreas, mouth, throat, pharynx, larynx or esophagus, or an upper airway primary or secondary cancer), to identify subjects who need an intensive follow-up protocol and for the prognosis and diagnosis of cancer. This is the amino acid sequence of the human ribosomal L14 protein (RPL14, located on chromosome 3p21.3), used to develop the gene probe described in the method of the invention.
Sequence 220 AA;
Query Match 100.0%; Score 45; DB 23; Length 220;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
1 AAAAAAAAAA 11
| | | | | | | | | |
155 AAAAAAAAAA 165
RESULT 13
ID ABOU7443
ABOU7443 standard; Protein; 220 AA.
ABOU7443;
28-JAN-2003 (first entry)
Protein differentially regulated in prostate cancer #46.
Prostate cancer; gene expression; differential regulation;
molecular marker; drug target; cancer detection; cancer diagnosis;
cancer staging; cancer grading; cancer assessing; cancer monitoring.
Homo sapiens.
WO200281638-A2.
17-OCT-2002.
08-APR-2002; 2002WO-US10824.
06-APR-2001; 2001US-281731P.
06-APR-2001; 2001US-281732P.

(ORIG-) ORIGENE TECHNOLOGIES INC.

PA Sun Z, Jay G;
XX WPI, 2003-058520/05.
DR

Novel genes which are differentially regulated in prostate cancer,
PT useful for diagnosing prostate cancer in prostate tissue sample and
PT assessing therapeutic or preventive intervention in prostate cancer
PT patients -
XX
XX Claim 1; Page 297; 416pp; English.
XX

The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity.
CC (I) is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is
CC useful in therapeutic applications to treat prostate cancer. The
CC identification of specific genes, and groups of genes, expressed in
CC pathways physiologically relevant to prostate cancer permits the
CC definition of functional and disease pathways and the delineation of
CC targets in these pathways which are useful in diagnostic, therapeutic,
CC and clinical applications. This is the amino acid sequence of a protein
CC differentially regulated in prostate cancer.
CC
CC Sequence 220 AA;
SQ

Query Match 100.0%; Score 45; DB 24; Length 220;
Best Local Similarity 100.0%; Prcd. No. 3.6;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 AAAAAAAAAAK 11
DB 155 AAAAAAAAAAK 155

RESULT 14
ABP62855
ID ABP62855 standard; Protein; 221 AA.
AC ABP62855;
XX
XX 14-OCT-2002 (first entry)
DT
DE Human polypeptide SEQ ID NO 292.
XX
XX Human; vulnerable; dermatological; neuroprotective; nootropic; cancer;
KM antiapoptosis; immunostimulant; cytotoxic; immunosuppressive;
KM antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KM burn; central nervous system disorder; Alzheimer's disease;

KW Parkinson's disease; Huntington's disease; immune disorder;
 KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
 OS Homo sapiens.
 XX WO200218424-A2.
 XX PD 07-MAR-2002.
 XX PF 31-AUG-2001; 2001WO-US27093.
 XX PR 01-SEP-2000; 2000US-0654935.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
 PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
 DR WPI; 2002-563321/62.
 DR N-PSDB; ABQ93334.
 XX PT New polynucleotide and polypeptides, useful for treatment and diagnosis
 PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
 PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
 PT sclerosis, diabetes and allergies -
 PS Claim 20; SEQ ID NO 292; 284pp + Sequence Listing; English.
 XX CC The invention relates to an isolated polynucleotide (I) comprising one of
 CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising
 CC administering to a mammalian subject a composition comprising the protein
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 221 AA;
 Query Match 100.0%; Score 45; DB 23; Length 221;
 Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAA 11
 DB 156 AAAAAAAAAA 166
 RESULT 15
 AAG00758
 ID AAG00758 standard; Protein: 225 AA.
 AC AAG00758;
 XX DT 06-OCT-2000 (first entry)
 XX DE Human secreted protein, SEQ ID NO: 4839.
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX OS Homo sapiens.
 XX EPI033401-A2.
 XX EN 06-SEP-2000.
 XX PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 XX DR N-PSDB; AAC00764.
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 13; SEQ ID 4839; 71pp + CD-ROM; English.
 XX CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX SQ Sequence 225 AA;
 Query Match 100.0%; Score 45; DB 21; Length 225;
 Best Local Similarity 100.0%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAA 11
 DB 162 AAAAAAAAAA 172
 Search completed: February 19, 2004, 10:03:18
 Job time : 42 secs

No data good

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 10:04:34 ; Search time 34 Seconds
(without alignments)
67.741 Million cell updates/sec

Title: US-09-551-336b-1
Perfect score: 45
Sequence: 1 AAAAAAAAAAK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	20	12	US-10-192-832-55 Sequence 55, Appl
2	45	100.0	30	12	US-10-192-832-68 Sequence 68, Appl
3	45	100.0	30	12	US-10-192-832-69 Sequence 69, Appl
4	45	100.0	79	12	US-10-177-725-22 Sequence 72, Appl
5	45	100.0	79	12	US-10-177-725-72 Sequence 72, Appl
6	45	100.0	97	15	US-10-106-698-6694 Sequence 6694, Ap
7	45	100.0	104	12	US-10-177-725-31 Sequence 31, Appl
8	45	100.0	104	12	US-10-177-725-32 Sequence 32, Appl
9	45	100.0	104	12	US-10-177-725-33 Sequence 33, Appl
10	45	100.0	104	12	US-10-177-725-34 Sequence 34, Appl
11	45	100.0	104	12	US-10-177-725-35 Sequence 35, Appl
12	45	100.0	104	12	US-10-177-725-36 Sequence 36, Appl
13	45	100.0	104	12	US-10-177-725-37 Sequence 37, Appl
14	45	100.0	104	12	US-10-177-725-38 Sequence 38, Appl
15	45	100.0	104	12	US-10-177-725-42 Sequence 42, Appl

16	45	100.0	104	12	US-10-177-725-81	Sequence 81, Appl
17	45	100.0	104	12	US-10-177-725-82	Sequence 82, Appl
18	45	100.0	104	12	US-10-177-725-83	Sequence 83, Appl
19	45	100.0	104	12	US-10-177-725-84	Sequence 84, Appl
20	45	100.0	104	12	US-10-177-725-85	Sequence 85, Appl
21	45	100.0	104	12	US-10-177-725-86	Sequence 86, Appl
22	45	100.0	104	12	US-10-177-725-87	Sequence 87, Appl
23	45	100.0	104	12	US-10-177-725-88	Sequence 88, Appl
24	45	100.0	104	12	US-10-177-725-92	Sequence 92, Appl
25	45	100.0	112	12	US-10-177-725-90	Sequence 90, Appl
26	45	100.0	112	12	US-10-177-725-80	Sequence 80, Appl
27	45	100.0	201	12	US-10-264-049-2597	Sequence 2597, Ap
28	45	100.0	220	9	US-09-923-304-2	Sequence 2, Appl
29	45	100.0	265	9	US-09-925-302-559	Sequence 559, Ap
30	45	100.0	617	12	US-10-104-047-2915	Sequence 2915, Ap
31	45	100.0	663	12	US-10-108-260A-2477	Sequence 2477, Ap
32	45	100.0	730	11	US-09-961-403-8	Sequence 8, Appl
33	45	100.0	731	12	US-09-964-662-1	Sequence 1, Appl
34	42	93.3	20	12	US-10-192-832-57	Sequence 57, Appl
35	42	93.3	30	12	US-10-192-832-70	Sequence 70, Appl
36	42	93.3	102	12	US-10-192-832-71	Sequence 9, Appl
37	42	93.3	102	12	US-10-121-016-9	Sequence 15, Appl
38	42	93.3	102	12	US-10-121-016-15	Sequence 70, Appl
39	42	93.3	102	12	US-10-121-016-70	Sequence 1325, Ap
40	42	93.3	157	10	US-09-764-864-1325	Sequence 2380, Ap
41	42	93.3	229	12	US-10-264-049-2380	Sequence 1902, Ap
42	42	93.3	276	10	US-09-902-941-1902	Sequence 1902, Ap
43	42	93.3	276	10	US-09-849-626-1902	Sequence 1902, Ap
44	42	93.3	276	12	US-10-113-872-1902	Sequence 1902, Ap
45	42	93.3	276	15	US-10-017-754-1902	Sequence 1902, Ap

ALIGNMENTS

RESULT 1
US-10-192-832-55
; Sequence 55, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, SHUANG
; APPLICANT: VAUTHEY, SYLVAIN
; TITLE OF INVENTION: SUBFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
; FILE REFERENCE: MTV-043.01
; CURRENT APPLICATION NUMBER: US/10/192,832
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/304,256
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-192-832-55

Query Match 100.0%; Score 45; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
Db 1 AAAAAAAAAAK 11

RESULT 2
US-10-192-832-68
; Sequence 68, Application US/10192832
; Publication No. US20030176335A1
; GENERAL INFORMATION:

```
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAUTHRY, SYLVAIN
TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTV-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 68
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-68
```

```
Query Match          100.0%; Score 45; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAAAAAK 11
    |||||
Db 11 AAAAAAAAAAK 21
```

```
RESULT 3
US-10-192-832-69
Sequence 69, Application US/10192832
Publication No. US20030176335A1
GENERAL INFORMATION:
APPLICANT: ZHANG, SHUGUANG
APPLICANT: VAUTHRY, SYLVAIN
TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF
FILE REFERENCE: MTV-043.01
CURRENT APPLICATION NUMBER: US/10/192,832
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/304,256
PRIOR FILING DATE: 2001-07-10
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 69
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-192-832-69
```

```
Query Match          100.0%; Score 45; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAAAAAK 11
    |||||
Db 1 AAAAAAAAAAK 11
```

```
RESULT 4
US-10-177-725-22
Sequence 22, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
```

```
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 79
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-177-725-22
```

```
Query Match          100.0%; Score 45; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAAAAAK 11
    |||||
Db 62 AAAAAAAAAAK 72
```

```
RESULT 5
US-10-177-725-72
Sequence 72, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72
LENGTH: 79
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
NAME/KEY: MISC FEATURE
LOCATION: (21)..(52)
OTHER INFORMATION: "Xaa" at positions 21-23, 25-27, 29-30, 32-34, 36-37, 39-41, 43-4
US-10-177-725-72
```

```
Query Match          100.0%; Score 45; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAAAAAK 11
    |||||
Db 62 AAAAAAAAAAK 72
```

```
RESULT 6
US-10-106-698-6694
Sequence 6694, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
FILE REFERENCE: PA005F1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
```

PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: Patentin Ver. 3.0
SEQ ID NO 6694
LENGTH: 97
TYPE: PRT
ORGANISM: Homo sapiens
US-10-106-698-6694

Query Match 100.0%; Score 45; DB 15; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
Db 77 AAAAAAAAAA 87

RESULT 7
US-10-177-725-31
Sequence 31, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: Patentin version 3.1
SEQ ID NO 31
LENGTH: 104
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-177-725-31

Query Match 100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
Db 93 AAAAAAAAAA 103

RESULT 8
US-10-177-725-32
Sequence 32, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015

PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: Patentin version 3.1
SEQ ID NO 32
LENGTH: 104
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-177-725-32

Query Match 100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
Db 93 AAAAAAAAAA 103

RESULT 9
US-10-177-725-33
Sequence 33, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: Patentin version 3.1
SEQ ID NO 33
LENGTH: 104
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-10-177-725-33

Query Match 100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
Db 92 AAAAAAAAAA 102

RESULT 10
US-10-177-725-34
Sequence 34, Application US/10177725
Publication No. US20030143562A1
GENERAL INFORMATION:
APPLICANT: Anderson, David
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Peele, Beau R.
TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT
FILE REFERENCE: A-66900-4/RMS/AMS
CURRENT APPLICATION NUMBER: US/10/177,725
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 173
SOFTWARE: Patentin version 3.1

```
; SEQ ID NO 34
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-34
```

```
Query Match          100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAAAAAAAAAK 11
        |||||
Db      82 AAAAAAAAAAK 92
```

```
RESULT 11
US-10-177-725-35
; Sequence 35, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-35
```

```
Query Match          100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAAAAAAAAAK 11
        |||||
Db      78 AAAAAAAAAAK 88
```

```
RESULT 12
US-10-177-725-36
; Sequence 36, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 104
; TYPE: PRT
```

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-36
```

```
Query Match          100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAAAAAAAAAK 11
        |||||
Db      76 AAAAAAAAAAK 86
```

```
RESULT 13
US-10-177-725-37
; Sequence 37, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-37
```

```
Query Match          100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 AAAAAAAAAAK 11
        |||||
Db      92 AAAAAAAAAAK 102
```

```
RESULT 14
US-10-177-725-38
; Sequence 38, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
```

US-10-177-725-38

Query Match 100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
|||
Db 92 AAAAAAAAAAK 102

RESULT 15

US-10-177-725-42

; Sequence 42, Application US/10177725
; Publication No. US20030143562A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David

; APPLICANT: Bogenberger, Jakob M.

; APPLICANT: Peele, Beau R.

; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S

; FILE REFERENCE: A-66900-4/RMS/AMS

; CURRENT APPLICATION NUMBER: US/10/177,725

; CURRENT FILING DATE: 2002-06-20

; PRIOR APPLICATION NUMBER: US 09/415,765

; PRIOR FILING DATE: 1999-10-08

; PRIOR APPLICATION NUMBER: US 09/169,015

; PRIOR FILING DATE: 1998-10-08

; NUMBER OF SEQ ID NOS: 173

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 42

; LENGTH: 104

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-10-177-725-42

Query Match 100.0%; Score 45; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
|||
Db 92 AAAAAAAAAAK 102

Search completed: February 19, 2004, 10:10:03
Job time : 35 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 19, 2004, 10:02:28 / Search time 22 Seconds
(without alignments)
21.155 Million cell updates/sec

Title: US-09-551-336B-1
Perfect score: 45
Sequence: 1 AAAAAAAAAAK 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:**

- 1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:**
- 2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:**
- 3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:**
- 4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:**
- 5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:**
- 6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	731	2	US-08-911-364-1
2	45	100.0	731	4	US-09-340-736E-1
3	45	100.0	733	3	US-08-464-700-2
4	45	100.0	792	2	US-08-678-039A-40
5	42	93.3	151	3	US-09-045-764A-4
6	41	91.1	39	4	US-09-117-121-28
7	41	91.1	54	4	US-09-117-121-30
8	41	91.1	238	4	US-09-334-332-10
9	41	91.1	238	4	US-09-702-705-1813
10	41	91.1	238	4	US-09-736-457-1813
11	41	91.1	340	3	US-09-360-779-2
12	41	91.1	340	4	US-09-435-335-2
13	41	91.1	11	3	US-09-208-966-53
14	40	88.9	12	1	US-07-392-288-7
15	40	88.9	12	1	US-07-989-764-7
16	40	88.9	12	3	US-09-058-459-37
17	40	88.9	12	3	US-09-127-926-37
18	40	88.9	12	3	US-08-788-822A-15
19	40	88.9	14	3	US-09-058-562-31
20	40	88.9	18	1	US-08-240-712-33
21	40	88.9	22	1	US-08-443-890-33
22	40	88.9	22	1	US-07-392-288-6
23	40	88.9	22	1	US-07-989-764-6
24	40	88.9	37	1	US-08-209-747-49
25	40	88.9	37	1	US-08-458-298-49
26	40	88.9	47	1	US-08-425-069-26
27	40	88.9	47	1	US-08-209-747-37

28	40	88.9	47	1	US-08-458-298-37	Sequence 37, Appl
29	40	88.9	47	2	US-08-317-844B-26	Sequence 26, Appl
30	40	88.9	47	3	US-08-556-978B-18	Sequence 18, Appl
31	40	88.9	177	3	US-09-058-562-27	Sequence 27, Appl
32	40	88.9	182	1	US-08-240-712-29	Sequence 29, Appl
33	40	88.9	182	1	US-08-443-890-29	Sequence 29, Appl
34	40	88.9	223	4	US-09-009-816-4	Sequence 4, Appl
35	40	88.9	223	4	US-09-140-749-51	Sequence 51, Appl
36	40	88.9	367	4	US-09-009-816-2	Sequence 2, Appl
37	40	88.9	407	2	US-08-765-875-2	Sequence 2, Appl
38	40	88.9	407	2	US-08-765-875-6	Sequence 6, Appl
39	40	88.9	407	3	US-08-795-671-2	Sequence 2, Appl
40	40	88.9	407	3	US-08-795-671-6	Sequence 6, Appl
41	40	88.9	407	4	US-09-454-540-2	Sequence 4, Appl
42	40	88.9	407	4	US-09-454-540-6	Sequence 6, Appl
43	40	88.9	442	3	US-09-347-833-11	Sequence 11, Appl
44	40	88.9	449	2	US-08-927-394-2	Sequence 2, Appl
45	40	88.9	453	6	5206152-7	Patent No. 5206152

ALIGNMENTS

RESULT 1
US-08-911-364-1
Sequence 1 Application US/08911364
Patent No. 5206106

GENERAL INFORMATION:

APPLICANT: ROTHSTEIN, Aser
APPLICANT: KEELY, Fred W.
TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
TITLE OF INVENTION: ELASTIN AND OTHER FIBROUS PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,364
FILING DATE: 07-AUG-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,552
FILING DATE: 07-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
TELEPHONE/DOCKET NUMBER: 041082/0104
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-911-364-1

8/7/97 Good

Query Match 100.0%, Score 45, DB 2, Length 731;
Best Local Similarity 100.0%, Pred. No. 1.8;
Matches 11, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
| | | | |
Db 276 AAAAAAAAAA 286

RESULT 2
US-09-340-736E-1
Sequence 1, Application US/09340736E
Patent No. 6489446
GENERAL INFORMATION:
APPLICANT: ROTHSTEIN, ASER
APPLICANT: KEELEY, FRED
APPLICANT: ROTHSTEIN, STEVEN
TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELLED ON HUMAN ELASTIN
TITLE OF INVENTION: AND OTHER FIBROUS PROTEINS
FILE REFERENCE: 041082/0110
CURRENT APPLICATION NUMBER: US/09/340,736E
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 08/911,364
PRIOR FILING DATE: 1997-08-07
PRIOR APPLICATION NUMBER: 60/023,552
PRIOR FILING DATE: 1996-08-07
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 731
TYPE: PRT
ORGANISM: Homo sapiens
US-09-340-736E-1

Query Match 100.0%; Score 45; DB 4; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
| | | | |
Db 276 AAAAAAAAAA 286

RESULT 3
US-08-678-700-2
Sequence 2, Application US/08464700
Patent No. 6232458
GENERAL INFORMATION:
APPLICANT: WEISS, ANTHONY S
APPLICANT: MARTIN, STEPHEN L
TITLE OF INVENTION: SYNTHETIC POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, PO Box 457
City: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,700
FILING DATE: 7-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL6520
FILING DATE: 22-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PL9661
FILING DATE: 28-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU93/00655
FILING DATE: 16-DEC-1993

ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GHCUUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 733 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-464-700-2

QY 1 AAAAAAAAAA 11
| | | | |
| | | | |
Db 278 AAAAAAAAAA 288

RESULT 4
US-08-678-039A-40
Sequence 40, Application US/08678039A
Patent No. 5858662
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
APPLICANT: Morris, Colleen A.
TITLE OF INVENTION: Diagnosis of Williams Syndrome and
TITLE OF INVENTION: Williams Syndrome Cognitive Profile by Analysis of the
TITLE OF INVENTION: Presence or Absence of a LIM-Kinase Gene
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurtz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701 East
STREET: Tower
City: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,039A
FILING DATE: 10-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Stephen A.
REGISTRATION NUMBER: 38,609
REFERENCE/DOCKET NUMBER: 2323-120A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-624-1589
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 792 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-678-039A-40

Query Match 100.0%; Score 45; DB 2; Length 792;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
| | | | |
| | | | |

Db 302 AAAAAAAAAAK 312

RESULT 5

US-09-045-764A-4
Sequence 4, Application US/09045764A
Patent No. 6127178

GENERAL INFORMATION:

APPLICANT: Israel, Mark A.

APPLICANT: Florio, Monica

TITLE OF INVENTION: Apoptotic Peptides

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/045,764A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UCSF98-045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 161 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-045-764A-4

Query Match 93.3%; Score 42; DB 3; Length 161;
Best Local Similarity 90.9%; Pred. No. 1.2;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
Db 39 AAAAAAAAAAK 49

RESULT 6

US-09-117-121-28

Sequence 28, Application US/09117121

Patent No. 6307020

GENERAL INFORMATION:

APPLICANT: Hew, Choy

APPLICANT: Gong, Zhiyuan

TITLE OF INVENTION: Intracellular Antifreeze Polypeptides

TITLE OF INVENTION: and Nucleic Acids

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/117,121

FILING DATE: 20-NOV-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/CA97/00062

FILING DATE: 30-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 016252-001610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-117-121-28

Query Match 91.1%; Score 41; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.45;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
Db 6 AAAAAATAAAK 16

RESULT 7

US-09-117-121-30

Sequence 30, Application US/09117121

Patent No. 6307020

GENERAL INFORMATION:

APPLICANT: Hew, Choy

APPLICANT: Gong, Zhiyuan

TITLE OF INVENTION: Intracellular Antifreeze Polypeptides

TITLE OF INVENTION: and Nucleic Acids

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/117,121

FILING DATE: 20-NOV-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/CA97/00062

FILING DATE: 30-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 016252-001610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 54 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-117-121-30

Query Match 91.1%; Score 41; DB 4; Length 54;
Best Local Similarity 90.9%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
Db 6 AAAAAATPAAAK 16

RESULT 8

US-09-234-332-10
Sequence 10, Application US/09234332A
Patent No. 6087168
GENERAL INFORMATION:
APPLICANT: Cedars-Sinai Medical Center
APPLICANT: Michel F. Levesque, M.D.
APPLICANT: Thomas Neuman, Ph.D.
TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
FILE REFERENCE: P07 41494
CURRENT APPLICATION NUMBER: US/09/234,332A
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 238
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (0)...(0)
OTHER INFORMATION: Achete scute homologous protein (ASH1); Genbank
US-09-234-332-10

Query Match 91.1%; Score 41; DB 3; Length 238;
Best Local Similarity 90.9%; Pred. No. 2.7;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
Db 38 AAAAAAAAAAQ 48

RESULT 9

US-09-702-705-1813
Sequence 1813, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1813
LENGTH: 238
TYPE: PRT
ORGANISM: Homo sapiens

US-09-702-705-1813

Query Match 91.1%; Score 41; DB 4; Length 238;
Best Local Similarity 90.9%; Pred. No. 2.7;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
Db 38 AAAAAAAAAAQ 48

RESULT 10

US-09-736-457-1813
Sequence 1813, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1813
LENGTH: 238
TYPE: PRT
ORGANISM: Homo sapiens
US-09-736-457-1813

Query Match 91.1%; Score 41; DB 4; Length 238;
Best Local Similarity 90.9%; Pred. No. 2.7;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
Db 38 AAAAAAAAAAQ 48

RESULT 11

US-09-360-779-2
Sequence 2, Application US/09360779
Patent No. 6268216
GENERAL INFORMATION:
APPLICANT: Deneris, Evan S.
APPLICANT: Eyodoro, Dmitry V.
APPLICANT: Hendricks, Timothy J.
TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
FILE REFERENCE: CASE-03828
CURRENT APPLICATION NUMBER: US/09/360,779
CURRENT FILING DATE: 1999-07-26
EARLIER APPLICATION NUMBER: 60/094,264
EARLIER FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 340
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-360-779-2

Query Match 91.1%; Score 41; DB 3; Length 340;
Best Local Similarity 90.9%; Pred. No. 3.9;

```
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 11
   |||||
   249 AAAAAAAAAA 259
Db

RESULT 12
US-09-435-335-2
; Sequence 2, Application US/09435335
; Patent No. 6384204
; GENERAL INFORMATION:
; APPLICANT: Denetis, Evan S.
; APPLICANT: Eyodoro, Dmitry V.
; APPLICANT: Hendricks, Timothy J.
; TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
; FILE REFERENCE: CASE-04027
; CURRENT APPLICATION NUMBER: US/09/435,335
; EARLIER FILING DATE: 1999-11-05
; EARLIER APPLICATION NUMBER: 09/360,779
; EARLIER FILING DATE: 1999-07-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-435-335-2

Query Match 91.1%; Score 41; DB 4; Length 340;
Best Local Similarity 90.9%; Pred. No. 3.9;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 11
   |||||
   249 AAAAAAAAAA 259
Db

RESULT 13
US-09-208-966-53
; Sequence 53, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48861/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 53
; LENGTH: 11
; TYPE: PRT
; ORGANISM: human
US-09-208-966-53

Query Match 88.9%; Score 40; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 10
   |||||
   2 AAAAAAAAAA 11
Db

RESULT 14
US-07-992-288-7
; Sequence 7, Application US/07992288
```

```
; Patent No. 5338831
; GENERAL INFORMATION:
; APPLICANT: Lebel, Michel
; APPLICANT: Richler, Jutta
; APPLICANT: Pokorny, Vite
; APPLICANT: Jehnicka, Jiri
; APPLICANT: Mudra, Petr
; APPLICANT: Zenisek, Karel
; APPLICANT: Stierandova, Alena
; APPLICANT: Kalousek, Jan
; APPLICANT: Bolf, Jan
; TITLE OF INVENTION: METHOD OF MAKING MULTIPLE SYNTHESIS OF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dressler, Goldsmith, Shore & Milnamow, Ltd.
; STREET: 180 No. 5338831th Stetson, Suite 4700
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/992,288
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/645,121
; FILING DATE: 24-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoover, Allen J.
; REGISTRATION NUMBER: 24,103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-992-288-7

Query Match 88.9%; Score 40; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 10
   |||||
   1 AAAAAAAAAA 10
Db

RESULT 15
US-07-989-764-7
; Sequence 7, Application US/07989764
; Patent No. 5342585
; GENERAL INFORMATION:
; APPLICANT: Lebel, Michel
; APPLICANT: Richler, Jutta
; APPLICANT: Pokorny, Vite
; APPLICANT: Jehnicka, Jiri
; APPLICANT: Mudra, Petr
; APPLICANT: Zenisek, Karel
; APPLICANT: Stierandova, Alena
; APPLICANT: Kalousek, Jan
; APPLICANT: Bolf, Jan
; TITLE OF INVENTION: APPARATUS FOR MAKING MULTIPLE SYNTHESIS
; TITLE OF INVENTION: OF PEPTIDES ON SOLID SUPPORT
; NUMBER OF SEQUENCES: 7
```

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dressler, Goldsmith, Shore & Milnamow, Ltd.
;; STREET: 180 No. 5342585th Stetson, Suite 4700
;; City: Chicago
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60601
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/989,764
;; FILING DATE:
;; CLASSIFICATION: 530
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/645,121
;; FILING DATE: 24-JAN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hoover, Allen J.
;; REGISTRATION NUMBER: 24,103
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312)616-5400
;; TELEFAX: (312)616-5460
;;
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 12 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
;; MOLECULE TYPE: peptide
;;
;; US-07-989-764-7

Query Match 88.9%; Score 40; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 10
|||
Db 1 AAAAAAAAAA 10

Search completed: February 19, 2004, 10:05:37
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 10:01:38 ; Search time 21 Seconds
(without alignments)
50.374 Million cell updates/sec

Title: US-09-551-336B-1
Perfect score: 45
Sequence: 1 AAAAAAAAAAK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
955 1	45	100.0	220	2 JC5954	ribosomal protein
2	45	100.0	435	2 A44308	Antho-RFamide precursor
3	45	100.0	792	1 EAHU	elastin precursor
4	45	100.0	1430	2 T34516	hypothetical prote
5	42	93.3	161	2 JEO306	Id4 protein - huma
6	42	93.3	161	2 G01855	helix-loop-helix p
7	42	93.3	161	2 S43260	unknown protein F9
8	42	93.3	333	2 G96780	hypothetical prote
9	42	93.3	814	2 T47641	serine-rich protei
10	42	93.3	1077	2 A44067	abdominal segment
11	42	93.3	1533	2 A46221	long chain fatty a
12	42	93.3	1607	2 T02837	fibroin - Chinese
13	42	93.3	2639	2 T31328	homoeotic protein u
14	41	91.1	40	2 S58853	antifreeze protein
15	41	91.1	97	2 S02376	achae-scute locu
16	41	91.1	231	2 S28186	probable MASH-2 pr
17	41	91.1	233	2 S11563	hypothetical prote
18	41	91.1	238	2 T11718	hypothetical prote
19	41	91.1	238	2 A46279	achae-scute prote
20	41	91.1	273	2 T51010	hypothetical prote
21	41	91.1	287	2 T51011	hypothetical prote
22	41	91.1	314	2 JC5273	paired type homeob
23	41	91.1	323	2 S16318	homoeotic protein H
24	41	91.1	337	2 S06956	segmentation prote
25	41	91.1	392	2 B46423	homoeotic protein e
26	41	91.1	451	1 A40168	transcription fact
27	41	91.1	451	1 D88395	protein F53A3.6 (l
28	41	91.1	649	2 S43229	arylsulfatase (EC
29	41	91.1	701	1 S46458	transcription fact

30	41	91.1	702	2 G01840	T-box protein 2 -
31	41	91.1	747	1 EABO	elastin precursor,
32	41	91.1	748	2 T49633	glucan 1,4-alpha-g
33	41	91.1	770	2 S59623	tropoelastin - she
34	41	91.1	796	2 UC7555	C1orf4 protein -
35	41	91.1	860	1 EAMS	elastin precursor
36	41	91.1	864	1 EART	elastin precursor
37	41	91.1	1065	2 T13230	dachshund isoform
38	41	91.1	1072	2 T13232	dachshund protein
39	41	91.1	1074	2 T13229	dachshund protein
40	41	91.1	1081	2 T13231	dachshund protein
41	41	91.1	1175	2 JH0697	potassium channel
42	41	91.1	1180	2 S69205	stripe a/b protein
43	41	91.1	1184	2 A39800	calcium-activated
44	41	91.1	1355	2 S40022	spalt protein - fr
45	41	91.1	1402	2 S42748	finger protein - f

ALIGNMENTS

RESULT 1

JC5954 ribosomal protein L14 - human
C:Species: Homo sapiens (man)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
C:Accession: JC5954
R:Tanaka, M.; Tanaka, T.; Harata, M.; Suzuki, T.; Mitsui, Y.
Biochem. Biophys. Res. Commun. 243, 531-537, 1998
A>Title: Triplet repeat-containing ribosomal protein L14 gene in immortalized human endo
A:Reference number: JC5954; MUID:98153799; PMID:9480843
A:Accession: JC5954
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-220 <TAN>
A:Cross-references: DDBJ:D87735; NID:G1620021; PID:BA113443.1; PID:G1620022
C:Superfamily: rat ribosomal protein L14

Query Match 100.0%; Score 45; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
DB 155 AAAAAAAAAAK 165

RESULT 2
A4308 Antho-RFamide precursor - sea anemone (Anthopleura elegantissima)

C:Species: Anthopleura elegantissima
C>Date: 10-Jun-1995 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
R:Schmutzler, C.; Darner, D.; Diekhoff, D.; Grimmelikhuijzen, C.J.
J. Biol. Chem. 267, 22534-22541, 1992
A>Title: Identification of a novel type of processing sites in the precursor for the
A:Reference number: A4308; MUID:93054550; PMID:1422603
A:Accession: A4308
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-435 <SCH>
A:Cross-references: GB:M98269; NID:G155702; PID:AAA27738.1; PID:G155703
A>Note: sequence extracted from NCBI backbone (NCBI:117102, NCBI:P:117104)
C:Keywords: neuropeptide

Query Match 100.0%; Score 45; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
DB 377 AAAAAAAAAAK 387

RESULT 3

A/HU

elastin precursor, long splice form - human

M/Alternate names: tropoelastin

C/Species: Homo sapiens (man)

C/Date: 22-Jun-1990 #sequence_revision 26-Jul-1996 #text_change 22-Jun-1999

C/Accession: A32707; A33705; A30524; A53891

R/Indik, Z.; Yeh, H.; Ornstein-Goldstein, N.; Sheppard, P.; Anderson, N.; Rosenbloom, J.

Proc. Natl. Acad. Sci. U.S.A. 84, 5680-5684, 1987

A/Title: Alternative splicing of human elastin mRNA indicated by sequence analysis of cDNA

A/Reference number: A32707; MUID:87289668; PMID:3039501

A/Accession: A32707

A/Molecule type: mRNA

A/Residues: 1-500,507-792 <IND>

A/Cross-references: GB:M16983; GB:J02948

R/Bashir, M.M.; Indik, Z.; Yeh, H.; Ornstein-Goldstein, N.; Rosenbloom, J.C.; Abrams, W.

J. Biol. Chem. 264, 8887-8891, 1989

A/Title: Characterization of the complete human elastin gene. Delineation of unusual features

A/Reference number: A33705; MUID:89255358; PMID:2722804

A/Accession: A33705

A/Molecule type: DNA

A/Residues: 1-27 <BAS>

A/Cross-references: GB:J04821; NID:g182052; PIDN:AAA52379.1; PID:g553276

R/Fazio, M.J.; Olsen, D.R.; Kaub, E.A.; Baldwin, C.T.; Indik, Z.; Ornstein-Goldstein, N.

J. Invest. Dermatol. 91, 458-464, 1988

A/Title: Cloning of full-length elastin cDNAs from a human skin fibroblast recombinant

A/Reference number: A30524; MUID:89009960; PMID:3171221

A/Accession: A30524

A/Molecule type: mRNA

A/Residues: 1-453,483-617,651-792 <FA2>

A/Cross-references: EMBL:M56860; NID:g182061; PIDN:AAA52382.1; PID:g182062

R/Fazio, M.J.; Olsen, D.R.; Kuivaniemi, H.; Chu, M.L.; Davidson, J.M.; Rosenbloom, J.

Lab. Invest. 58, 270-277, 1988

A/Title: Isolation and characterization of human elastin cDNAs, and age-associated variation

A/Reference number: A53891; MUID:88156138; PMID:2831431

A/Accession: A53891

A/Molecule type: mRNA

A/Residues: 164-453,483-500,507-617,651-792 <FA2>

A/Cross-references: GB:M24782; NID:g182063; PIDN:AAA53190.1; PID:g182064

C/Comment: The term tropoelastin refers to a soluble precursor form of the extracellular

C/Genetics:

A/Status: preliminary

A/Map position: 7q11.23-7q11.23

A/Cross-references: GDB:119107; OMIM:130160

C/Superfamily: elastin

C/Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine

F/1-26/Domain: signal sequence #status predicted <SIG>

F/72-792/Product: elastin #status predicted <MAT>

F/782-787/Disulfide bonds: #status predicted

Query Match 100.0%; Score 45; DB 1; Length 792;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11

Db 302 AAAAAAAAAA 312

RESULT 4

T34516

hypothetical protein ZK783.4 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Sep-2000

C/Accession: Z34516

R/Favella, A.; Vaudin, M.

Submitted to the EMBL Data Library, August 1994

Description: The sequence of C. elegans cosmid ZK783.

Reference number: Z21536

A/Accession: T34516

A/Status: preliminary; translated from GB/EMBL/DDBU

A/Molecule type: DNA

A/Residues: 1-1430 <FAV>

A/Cross-references: EMBL:U13646; PIDN:AAC24421.1; GSPDB:GN00021; CESP:ZK783.4

A/Experimental source: strain Bristol N2; clone ZK783

C/Genetics:

A/Status: preliminary

A/Map position: 3

A/Intons: 248/3; 373/3; 547/1; 593/2; 905/2; 1042/3; 1116/3; 1317/2; 1376/1

C/Superfamily: bromodomain homology

F/1284-1339/Domain: bromodomain homology <BRO>

Query Match 100.0%; Score 45; DB 2; Length 1430;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11

Db 114 AAAAAAAAAA 124

RESULT 5

JB0306

Id4 protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000

C/Accession: JB0306

R/Rigout, M.; Rich, T.; Gross-Morand, M.; Molina-Gomes, D.; Viegas-Pequignot, E.; Junn

DNA Res. 5, 309-313, 1998

A/Title: CDNA cloning, tissue distribution and chromosomal localization of the human ID4

A/Reference number: JB0306; MUID:99087490; PMID:9872455

A/Accession: JB0306

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-161 <RIG>

A/Cross-references: GB:Y07958

C/Superfamily: transcription repressor Id-2

Query Match 93.3%; Score 42; DB 2; Length 161;

Best Local Similarity 90.9%; Pred. No. 15;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11

Db 39 AAAAAAAAAA 49

RESULT 6

G01855

helix-loop-helix protein Id4 - human

C/Species: Homo sapiens (man)

C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 19-May-2000

C/Accession: G01855

R/Kiesling, T.L.

Submitted to the EMBL Data Library, June 1995

A/Reference number: G08632

A/Accession: G01855

A/Status: preliminary; translated from GB/EMBL/DDBU

A/Molecule type: mRNA

A/Residues: 1-161 <RLE>

A/Cross-references: EMBL:U28368; NID:g881545; PIDN:AAA73923.1; PID:g881546

C/Superfamily: transcription repressor Id-2

Query Match 93.3%; Score 42; DB 2; Length 161;

Best Local Similarity 90.9%; Pred. No. 15;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11

Db 39 AAAAAAAAAA 49

RESULT 7

S43260
 helix-loop-helix protein Id4, dominant negative - mouse
 A:Species: Mus musculus (house mouse)
 C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999
 C:Accession: S43260; S37199
 R:Riechmann, V.; van Crecchten, I.; Sablitzky, F.
 Nucleic Acids Res. 22, 749-755, 1994
 A:Title: The expression pattern of Id4, a novel dominant negative helix-loop-helix protein
 A:Reference number: S43260; MUID:94188125; PMID:8139914
 A:Accession: S43260
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-161 <R1E>
 A:Cross-references: EMBL:X75018; NID:9402637; PIDN:CAA52926.1; PID:9402638
 C:Superfamily: transcription repressor Id-2
 F:64-106/Region: helix-loop-helix #stratus predicted

Query Match 93.3%; Score 42; DB 2; Length 161;
 Best Local Similarity 90.9%; Pred. No. 15;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 39 AAAAAAAAAAK 49

RESULT 8

G96780
 unknown protein P9E10.7 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: G96780
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Coml, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Lin, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luero, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G96780
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <STO>
 A:Cross-references: GB:AE005173; NID:g6646757; PIDN:AAF21069.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: P9E10.7
 A:Map position: 1

Query Match 93.3%; Score 42; DB 2; Length 333;
 Best Local Similarity 90.9%; Pred. No. 25;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 12 AAAAAAAAAAK 22

RESULT 9

T47641
 hypothetical protein T15C9.20 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47641
 R:Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24470
 A:Accession: T47641
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-814 <MEM>
 A:Cross-references: EMBL:AL132970
 A:Experimental source: cultivar Columbia; BAC clone T15C9
 C:Genetics:
 A:Map position: 3
 A:Introns: 28/2; 51/3; 252/3; 312/3; 350/3; 370/2; 392/3; 410/1; 441/1; 469/1; 604/3; 62
 A:Note: T15C9.20

Query Match 93.3%; Score 42; DB 2; Length 814;
 Best Local Similarity 90.9%; Pred. No. 49;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 6 AAAAAAAAAAK 16

RESULT 10

A44067
 serine-rich protein hairless - fruit fly (Drosophila melanogaster)
 N:Alternate names: 109K basic protein H
 C:Species: Drosophila melanogaster
 C:Date: 10-Jun-1993 #sequence_revision 26-Feb-1999 #text_change 26-Feb-1999
 C:Accession: A44067; A58929; S33412; S24639
 R:Bang, A.G.; Posakony, J.W.
 Genes Dev. 6, 1752-1769, 1992
 A:Title: The Drosophila gene Hairless encodes a novel basic protein that controls altern
 A:Reference number: A44067; MUID:92387549; PMID:1516831
 A:Accession: A44067
 A:Molecule type: DNA
 A:Residues: 19-1077 <BAN>
 A:Cross-references: GB:M95192; NID:9157621; PID:9157622
 A:Note: sequence extracted from NCBI backbone (NCBIN:112622, NCBI:112623)
 R:Preiss, A.
 submitted to the EMBL Data Library, May 1994
 A:Description: Hairless, a Drosophila gene involved in neural development, encodes a nov
 A:Reference number: A58929
 A:Accession: A58929
 A:Molecule type: mRNA
 A:Residues: 1-1077 <PRE>
 A:Cross-references: EMBL:X67239; GB:S49642; NID:9578331; PID:9578332
 R:Maier, D.; Stumm, G.; Kuhn, K.; Preiss, A.
 Mech. Dev. 38, 143-156, 1992
 A:Title: Hairless, a Drosophila gene involved in neural development, encodes a novel, se
 A:Reference number: S33412; MUID:93041287; PMID:1419850
 A:Accession: S33412
 A:Molecule type: mRNA
 A:Residues: 1-150, 'A', '152-701', '704-890', 'R', '892-963', 'RLIP', '968-973', '975-1077 <MAI>
 A:Cross-references: EMBL:X67239
 C:Genetics:
 A:Gene: FlyBase:H; hairless
 A:Cross-references: FlyBase:FBgn0001169

Query Match 93.3%; Score 42; DB 2; Length 1077;
 Best Local Similarity 90.9%; Pred. No. 61;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 937 AAAAAAAAAAK 947

RESULT 11

A46221
 abdominal segment formation protein puntllo - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 21-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 17-Oct-1997
 C:Accession: A46221; S22026
 R:Barker, D.D.; Wang, C.; Moore, J.; Dickinson, L.K.; Lehmann, R.
 Genes Dev. 6, 2312-2326, 1992
 A:Title: Puntllo is essential for function but not for distribution of the Drosophila at
 A:Reference number: A46221; MUID:93093466; PMID:1459455

```

A:Accession: A46221
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1533 <BAR>
A:Cross-references: GB:107943; NID:g158190; PID:g158191
A>Note: sequence extracted from NCBI backbone (NCBIN:120203, NCBIPI:120204)
R:Macdonald, P.M.
submitted to the EMBL Data Library, October 1991
A:Reference number: S22026
A:Accession: S22026
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-361, 'A', '363-1102', 'R', '1104-1405', 'KN', '1408-1495', 'V', '1497-1518', 'S', '1520-1533
A:Cross-references: EMBL:X62589; NID:g83393; PID:g8394
C:Genetic8:
A:Gene: FLYBase:pum
A:Cross-references: FLYBase:FBgn0003165

Query Match          93.3%; Score 42; DB 2; Length 1533;
Best Local Similarity 90.9%; Pred. No. 79;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAAAAAAAAK 11
        |||||:
DB      937 AAAAAAAAAAAR 947

RESULT 12
T02837
long chain fatty acyl CoA synthetase LCFACAS5 [imported] - Leishmania major (strain Friedl
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
A:Accession: A81461; 102837
R:Myler, P.J.; Audleman, L.; d'Avos, T.; Hixson, G.; Kiber, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: A81461
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1607 <PVL>
A:Cross-references: GB:AE001274; NID:G3264850; PIDN:AAC24660.1; PID:g1617560; GSPDB:GN00
C:Genetic8:
A:Experimental source: strain MHOM/IL/81/Friedlin
A:Gene: LCFACAS5
A:Map position: 1

Query Match          93.3%; Score 42; DB 2; Length 1607;
Best Local Similarity 90.9%; Pred. No. 82;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAAAAAAAAAAK 11
        |||||:
DB      674 AAAAAAAAAAAR 684

RESULT 13
T31328
fibroin - Chinese oak silkworm
C:Species: Antherea pernyi (Chinese oak silkworm)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
A:Accession: T31328
R:Sezutsu, H.; Tamura, T.; Yukuhiro, K.
submitted to the EMBL Data Library, August 1998
A:Description: Characterization of the full length fibroin gene of a wild silkworm, Anth
A:Reference number: Z20995
A:Accession: T31328
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2639 <SEZ>
A:Cross-references: EMBL:AF083334; NID:g3450882; PID:g3450883; PIDN:AAC32606.1
C:Genetic8:
A:introns: 14/3

```

```

Query Match      93.3%; Score 42; DB 2; Length 2639;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAAAAAAAAA 11
        |||||
Db       366 AAAAAAAAAA 376

```

RESULT 14

homeotic protein; ultrabithorax homolog - Junonia coenia (fragment)
N.Alternate names: ultrabithorax homeodomain protein
C.Species: Junonia coenia
C.Date: 19-Mar-1997 #sequence, revision 29-Aug-1997 #text, change 15-Oct-1999
C.Accession: S58853
R.Warren, R.W., Nagy, L.; Selegue, J.; Gates, J.; Carroll, S.
Nature 372, 458-461, 1994
A.Title: Evolution of homeotic gene regulation and function in flies and butterflies.
A.Reference number: S58850, MUID:95075455, PMID:7840822
A.Accession: S58853
A.Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: mRNA
A.Residues: 1-40 <VAR>
A.Cross-references: EMBL:J42137
A.Note: The nucleotide sequence was submitted to the EMBL Data Library, May 1995
C.Superfamily: unassigned homeobox proteins; homeobox data
C.Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query M

Best Local Similarity 90.9%; Pred. No. 7.2;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0

```
QY      1 AAAAAAAAAAK 11
         |||||:
Db      25 AAAAAAAAAAQ 35
```

RESULT 15

antifreeze protein precursor - yellowtail flounder
C/Species: limanda ferruginea (yellowtail flounder)
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 24-Oct-2000
C/Accession: S02376
R/Scott, G.K.; Davies, P.L.; Shears, M.A.; Fletcher, G.L.
Burr, J. Biochem. 168, 629-633, 1987
A/Title: Structural variations in the alanine-rich antifreeze proteins of the Pleuronectes
A/Reference number: S02376; PMID:88029483; PMID:3665937
A/Accession: S02376
A/Molecule type: mRNA
A/Residues: 1-97 <SC0>
A/Cross-references: EMBL:X06356; NID:G64041; PIDD:CAA29655.1; PID:G64042
A/Note: part of this sequence, including the amino end of the mature protein, was confi.
C/Superfamily: antifreeze protein
C/Keywords: antifreeze
F/1-23/Domain: signal sequence #status predicted <IG>
F/24-48/Domain: propeptide #status predicted <PO>
F/49-96/Product: antifreeze protein #status predicted <MT>

Query Match	91.1%;	Score 41;	DB 2;	Length 97;
Best Local Similarity	90.9%;	Pred. No. 14;		
Matches 10;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY 1 AAAAAAAAAAK 11
||| ||| |||
Db 57 AAAAAATAAAAAAK 67

Search completed: February 19, 2004, 10:05:03
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 10:01:13 / Search time 10 Seconds

(without alignments)
51.729 Million cell updates/sec

Title: US-09-551-336B-1
Perfect score: 45
Sequence: 1 AAAAAAAAAA 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB-seq length: 0
Maximum DB-seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	406	1 LHX2_HUMAN	P50458 homo sapien
2	45	100.0	406	1 LHX2_MOUSE	Q92082 mus musculu
3	45	100.0	435	1 ELM1_MOUSE	P10419 anthopleura
4	45	100.0	730	1 ELS_HUMAN	P15602 homo sapien
5	42	93.3	161	1 ID4_HUMAN	P47928 homo sapien
6	42	93.3	161	1 ID4_MOUSE	P41139 mus musculu
7	42	93.3	276	1 SX21_HUMAN	Q9Y651 mus musculu
8	42	93.3	280	1 SX21_CHICK	Q9W795 gallus gall
9	42	93.3	530	1 ZIC2_MOUSE	Q62520 mus musculu
10	42	93.3	764	1 CSM_DROVI	Q24708 drosophila
11	42	93.3	1077	1 HLB5_DROME	Q02308 drosophila
12	42	93.3	1533	1 PUM_DROME	P25822 drosophila
13	41	91.1	97	1 ANP_LIMFE	P09031 limanda fer
14	41	91.1	231	1 ASH1_MOUSE	Q02067 mus musculu
15	41	91.1	233	1 ASH1_RAT	P19359 rattus norv
16	41	91.1	236	1 ASH1_HUMAN	P50553 homo sapien
17	41	91.1	314	1 PMXB_HUMAN	Q99453 homo sapien
18	41	91.1	314	1 PMXB_MOUSE	Q35690 mus musculu
19	41	91.1	323	1 HXDB_MOUSE	P23133 mus musculu
20	41	91.1	337	1 HAIR_DROME	P13177 homo sapien
21	41	91.1	338	1 HXDB_HUMAN	Q05925 homo sapien
22	41	91.1	392	1 HMR1_HUMAN	P56224 brachydanto
23	41	91.1	437	1 ZP12_BRARE	Q90436 brachydanto
24	41	91.1	441	1 P033_BRARE	P20267 rattus norv
25	41	91.1	451	1 P031_RAT	P81067 mus musculu
26	41	91.1	507	1 IRX3_MOUSE	Q10723 volvox carl
27	41	91.1	701	1 ARS_VOLCA	Q60707 mus musculu
28	41	91.1	701	1 TBX2_MOUSE	Q13207 homo sapien
29	41	91.1	702	1 TBX2_HUMAN	P04385 bos taurus
30	41	91.1	747	1 ELS_BOVIN	P54320 mus musculu
31	41	91.1	860	1 ELS_MOUSE	Q99372 rattus norv
32	41	91.1	864	1 ELS_RAT	Q24206 drosophila
33	41	91.1	880	1 BRCT_DROME	

34	41	91.1	1067	1 BAB2_DROME	Q9W0K4 drosophila
35	41	91.1	1355	1 SALM_DROME	P39770 drosophila
36	41	91.1	1402	1 SALM_DROVI	P39806 drosophila
37	41	91.1	1596	1 MAM_DROME	P21519 drosophila
38	41	91.1	1669	1 ASX_DROME	Q9V727 drosophila
39	41	91.1	1783	1 RAH3_CHLRE	Q9TEC4 chlamydomon
40	40	88.9	109	1 RLAI_TRYCR	P26643 trypanosoma
41	40	88.9	217	1 HAN2_HUMAN	Q95300 homo sapien
42	40	88.9	289	1 HXDB_MOUSE	P23463 mus musculu
43	40	88.9	297	1 MBL_DROME	O16011 drosophila
44	40	88.9	304	1 GSH2_HUMAN	Q9BZM3 homo sapien
45	40	88.9	305	1 GSH2_MOUSE	P31316 mus musculu

ALIGNMENTS

RESULT 1	ID	Sequence	STANDARD	PRT	406 AA
LHX2_HUMAN	AC	P50458; Q95860;			
LHX2_HUMAN	DT	01-OCT-1996 (Rel. 34, Created)			
LHX2_HUMAN	DT	30-MAY-2000 (Rel. 39, Last sequence update)			
LHX2_HUMAN	DT	28-FEB-2003 (Rel. 41, Last annotation update)			
LHX2_HUMAN	DE	LIM/homeobox protein Lhx2 (Homeobox protein Lhx2)			
LHX2_HUMAN	GN	LHX2 OR LHX2			
LHX2_HUMAN	OS	Homo sapiens (Human)			
LHX2_HUMAN	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
LHX2_HUMAN	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
LHX2_HUMAN	OX	MGI Taxid:9606;			
LHX2_HUMAN	RP	SEQUENCE FROM N.A.			
LHX2_HUMAN	RC	TISSUE=Placenta;			
LHX2_HUMAN	RX	MEDLINE=96226351; PubMed=8649822;			
LHX2_HUMAN	RA	Wu H.-K., Heng H.H.Q., Siderovski D.P., Dong W.-F., Okuno Y.,			
LHX2_HUMAN	RA	Shi X.-M., Tsui L.-C., Minden M.D.,			
LHX2_HUMAN	RT	"Identification of a human LIM-Hox gene, hLH-2, aberrantly expressed			
LHX2_HUMAN	RL	in chronic myelogenous leukaemia and located on 9q33-34.1,"			
LHX2_HUMAN	RN	entogene 12:1205-1212 (1996).			
LHX2_HUMAN	RP	SEQUENCE OF 18-406 FROM N.A.			
LHX2_HUMAN	RX	MEDLINE=99162575; PubMed=10051612;			
LHX2_HUMAN	RA	Rincon-Limas D.E., Lu C.-H., Calleja M.,			
LHX2_HUMAN	RT	Rodriguez-Esteban C., Izpisua-Belmonte J.C., Botas J.			
LHX2_HUMAN	RL	"Conservation of the expression and function of apertous orthologs in			
LHX2_HUMAN	CC	Drosophila and mammals,"			
LHX2_HUMAN	CC	Proc. Natl. Acad. Sci. U.S.A. 96:2165-2170 (1999).			
LHX2_HUMAN	CC	-1- FUNCTION: TRANSCRIPTIONAL REGULATORY PROTEIN INVOLVED IN THE			
LHX2_HUMAN	CC	CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND			
LHX2_HUMAN	CC	NEURAL CELL TYPES (BY SIMILARITY).			
LHX2_HUMAN	CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
LHX2_HUMAN	CC	-1- SIMILARITY: Contains 1 homeobox domain.			
LHX2_HUMAN	CC	-1- SIMILARITY: Contains 2 LIM zinc-binding domains.			
LHX2_HUMAN	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
LHX2_HUMAN	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
LHX2_HUMAN	CC	the European Bioinformatics Institute. There are no restrictions on its			
LHX2_HUMAN	CC	use by non-profit institutions as long as its content is in no way			
LHX2_HUMAN	CC	modified and this statement is not removed. Usage by and for commercial			
LHX2_HUMAN	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
LHX2_HUMAN	CC	or send an email to license@sib-sib.ch).			
LHX2_HUMAN	CC	-----			
LHX2_HUMAN	CC	EMBL: U11701; AAB08752.1; -			
LHX2_HUMAN	CC	EMBL: AF124735; AAD20013.1; -			
LHX2_HUMAN	CC	HSSP: P06601; 1FJL.			
LHX2_HUMAN	CC	GeneW: HGNC:6594; LHX2.			
LHX2_HUMAN	CC	MIM: 603759; -			
LHX2_HUMAN	CC	GO: 0007048; P: oncogenesis; TAS.			
LHX2_HUMAN	CC	InterPro: IPR001356; Homeobox.			
LHX2_HUMAN	CC	InterPro: IPR001781; LIM.			
LHX2_HUMAN	CC	InterPro: IPR007107; LIM_homo.			
LHX2_HUMAN	CC	Pfam: PF00046; homeobox; 1.			
LHX2_HUMAN	CC	Pfam: PF00412; LIM; 2.			

DR Prodom; PD000010; Homeobox; 1.
DR Prodom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00478; LIM DOMAIN 1; 2.
DR PROSITE; PS50023; LIM DOMAIN 2; 2.
DR PROSITE; PS50027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KM Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
KM Metal-binding; Zinc; Transcription regulation.
FT DOMAIN 53 105 LIM 1.
FT DOMAIN 115 168 LIM 2.
FT DOMAIN 187 196 POLY-ALA.
FT DNA BIND 266 325 HOMEBOX.
FT DOMAIN 307 323 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT CONFLICT 22 28 AKSEAP -> QERGR (IN REF. 1).
FT CONFLICT 54 54 A -> Q (IN REF. 1).
FT CONFLICT 187 187 A -> G (IN REF. 1).
FT CONFLICT 190 192 AAA -> RAR (IN REF. 1).
FT CONFLICT 196 196 MISSING (IN REF. 1).
FT CONFLICT 241 242 NA -> TR (IN REF. 1).
FT CONFLICT 391 406 GHEPHSPQTLINLF -> AMSLTAHKKLLPTFSNDSP
PHTSLKKSLV (IN REF. 1).
SQ SEQUENCE 406 AA; 44373 MW; DF73AAC81867D30D CRC64;
Query Match 100.0%; Score 45; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 4.6; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 11
Db 187 AAAAAAAAAA 197
RESULT 2
LHX2 MOUSE STANDARD; PRT; 406 AA.
ID LHX2 MOUSE STANDARD; PRT; 406 AA.
AC 092052;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE LIM/homeobox protein lhx2.
GN LHX2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=99162575; PubMed=10051612;
RA Rincon-Limas D.E., Lu C.-H., Canal I., Calleja M.,
RA Rodriguez-Besteiro C., Izpisua-Belmonte J.C., Botos J.;
RT "Conservation of the expression and function of apterous orthologs in
RT Drosophila and mammals.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2165-2170(1999).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR PROTEIN INVOLVED IN THE
CC CONTROL OF CELL DIFFERENTIATION IN DEVELOPING LYMPHOID AND
CC NEURAL CELL TYPES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 homeobox domain.
CC -1- SIMILARITY: Contains 2 LIM zinc-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF124734; AAD20012.1; -
CC HSSP; P06601; 1FJL.
CC TRANSFAC; T01969; -

DR MGD; MGI:96785; Lhx2.
DR GO; GO:0007498; P:neuroderm development; IMP.
DR GO; GO:0007399; P:neurogenesis; IMP.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR001781; LIM.
DR InterPro; IPR007107; LIM_homeo.
DR Pfam; PF00412; LIM; 2.
DR Pfam; PF00046; homeobox; 1.
DR Prodom; PD000010; Homeobox; 1.
DR Prodom; PD000094; LIM; 2.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00132; LIM; 2.
DR PROSITE; PS00478; LIM DOMAIN 1; 2.
DR PROSITE; PS50023; LIM DOMAIN 2; 2.
DR PROSITE; PS50027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KM Homeobox; DNA-binding; Nuclear protein; Repeat; LIM domain;
KM Metal-binding; Zinc; Transcription regulation.
FT DOMAIN 53 105 LIM 1.
FT DOMAIN 115 168 LIM 2.
FT DOMAIN 187 196 POLY-ALA.
FT DNA BIND 266 325 HOMEBOX.
FT DOMAIN 307 323 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 406 AA; 44419 MW; FE7B4E76454D6A90 CRC64;
Query Match 100.0%; Score 45; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 4.6; Mismatches 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAAAAA 11
Db 187 AAAAAAAAAA 197
RESULT 3
FMR1 ANTEL STANDARD; PRT; 435 AA.
ID FMR1 ANTEL STANDARD; PRT; 435 AA.
AC P10419;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Antho-Rfamde neuropeptides type 1 precursor.
OS Anthopleura elegantissima (sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=93054550; PubMed=1429603;
RA Schmutzler C., Darmer D., Diekhoff D., Grimmelikhuijzen C.J.P.;
RT "Identification of a novel type of processing sites in the precursor
RT for the sea anemone neuropeptide Antho-Rfamde (<Glu-Gly-Arg-Phe-NH2)
RT from Anthopleura elegantissima.";
RL J. Biol. Chem. 267:22534-22541 (1992).
RN (2)
RP PARTIAL SEQUENCE (ANTHO-RFAMIDE).
RX MEDLINE=87092339; PubMed=2879288;
RA Grimmelikhuijzen C.J.P., Graff D.;
RT "Isolation of Pyroglu-Gly-Arg-Phe-NH2 (Antho-Rfamide), a neuropeptide
RT from sea anemones.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9817-9821(1986).
CC -1- FUNCTION: NOT KNOWN BUT IT COULD ACT AS A TRANSMITTER AT
CC NEURONISCUAR SYNAPSES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M98269; AAA27738.1; -.
 DR PIR; A44308; A44308.
 DR InterPro; IPR002544; FARP.
 DR Pfam; PF01581; FARP; 13.
 DR Neuropeptide; Amidation; Repeat; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 22
 FT PEPTIDE 194 197 ANTHO-RFAMIDE.
 FT PEPTIDE 202 205 ANTHO-RFAMIDE.
 FT PEPTIDE 210 213 ANTHO-RFAMIDE.
 FT PEPTIDE 218 221 ANTHO-RFAMIDE.
 FT PEPTIDE 226 229 ANTHO-RFAMIDE.
 FT PEPTIDE 234 237 ANTHO-RFAMIDE.
 FT PEPTIDE 242 245 ANTHO-RFAMIDE.
 FT PEPTIDE 250 253 ANTHO-RFAMIDE.
 FT PEPTIDE 258 261 ANTHO-RFAMIDE.
 FT PEPTIDE 266 269 ANTHO-RFAMIDE.
 FT PEPTIDE 274 277 ANTHO-RFAMIDE.
 FT PEPTIDE 282 285 ANTHO-RFAMIDE.
 FT PEPTIDE 290 293 ANTHO-RFAMIDE.
 FT PEPTIDE 298 301 ANTHO-RFAMIDE.
 FT PEPTIDE 306 309 ANTHO-RFAMIDE.
 FT PEPTIDE 314 317 ANTHO-RFAMIDE.
 FT PEPTIDE 322 325 ANTHO-RFAMIDE.
 FT PEPTIDE 330 333 ANTHO-RFAMIDE.
 FT PEPTIDE 343 346 ANTHO-RFAMIDE.
 FT PEPTIDE 356 359 ANTHO-RFAMIDE.
 FT PEPTIDE 369 372 ANTHO-RFAMIDE.
 FT DOMAIN 376 386 POLY-ALA.
 FT MOD_RES 194 194 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 197 197 AMIDATION (G-198 PROVIDE AMIDE GROUP).
 FT MOD_RES 205 205 AMIDATION (G-206 PROVIDE AMIDE GROUP).
 FT MOD_RES 213 213 AMIDATION (G-214 PROVIDE AMIDE GROUP).
 FT MOD_RES 221 221 AMIDATION (G-222 PROVIDE AMIDE GROUP).
 FT MOD_RES 229 229 AMIDATION (G-230 PROVIDE AMIDE GROUP).
 FT MOD_RES 237 237 AMIDATION (G-238 PROVIDE AMIDE GROUP).
 FT MOD_RES 245 245 AMIDATION (G-246 PROVIDE AMIDE GROUP).
 FT MOD_RES 253 253 AMIDATION (G-254 PROVIDE AMIDE GROUP).
 FT MOD_RES 261 261 AMIDATION (G-262 PROVIDE AMIDE GROUP).
 FT MOD_RES 269 269 AMIDATION (G-270 PROVIDE AMIDE GROUP).
 FT MOD_RES 277 277 AMIDATION (G-278 PROVIDE AMIDE GROUP).
 FT MOD_RES 285 285 AMIDATION (G-286 PROVIDE AMIDE GROUP).
 FT MOD_RES 293 293 AMIDATION (G-294 PROVIDE AMIDE GROUP).
 FT MOD_RES 301 301 AMIDATION (G-302 PROVIDE AMIDE GROUP).
 FT MOD_RES 309 309 AMIDATION (G-310 PROVIDE AMIDE GROUP).
 FT MOD_RES 317 317 AMIDATION (G-318 PROVIDE AMIDE GROUP).
 FT MOD_RES 325 325 AMIDATION (G-326 PROVIDE AMIDE GROUP).
 FT MOD_RES 333 333 AMIDATION (G-334 PROVIDE AMIDE GROUP).
 FT MOD_RES 346 346 AMIDATION (G-347 PROVIDE AMIDE GROUP).
 FT MOD_RES 359 359 AMIDATION (G-360 PROVIDE AMIDE GROUP).
 FT MOD_RES 372 372 AMIDATION (G-373 PROVIDE AMIDE GROUP).
 SO SEQUENCE 435 AA; 50940 MW; B0C44020CD5861 CRC64;

Query Match 100.0%; Score 45; DB 1; Length 435;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
 DB 377 AAAAAAAAAA 387

RESULT 4
 ELS_HUMAN STANDARD; PRT; 730 AA.
 AC P15502; Q14233; Q14238;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Elastin precursor (Tropoelastin).
 GN ELN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RX MEDLINE=87289668; PubMed=3039501;
 RA Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Pelton L., Rosenbloom J.;
 RA "Alternative splicing of human elastin mRNA indicated by sequence
 analysis of cloned genomic and complementary DNA."
 Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Skin fibroblast;
 RX MEDLINE=89009960; PubMed=3171221;
 RA Fazio M.J., Olsen D.R., Kaub E.A., Baldwin C.T., Indik Z.,
 RA Ornstein-Goldstein N., Yeh H., Rosenbloom J., Ulitto J.;
 RA "Cloning of full-length elastin cDNAs from a human skin fibroblast
 recombinant cDNA library: further elucidation of alternative splicing
 utilizing exon-specific oligonucleotides."
 Invest. Dermatol. 91:458-464(1988).
 RN (3)
 RP SEQUENCE OF 164-724 FROM N.A. (ISOFORM B).
 RC TISSUE=Placenta;
 RX MEDLINE=88156138; PubMed=2831431;
 RA Fazio M.J., Olsen D.R., Kuivaniemi H., Chu M.L., Davidson J.M.,
 RA Rosenbloom J., Ulitto J.;
 RA "Isolation and characterization of human elastin cDNAs, and age-
 associated variation in elastin gene expression in cultured skin
 fibroblasts."
 Invest. 58:270-277(1988).
 RN (4)
 RP SEQUENCE OF 603-730 FROM N.A.
 RC TISSUE=Hippocampus, and Placenta;
 RX MEDLINE=96291399; PubMed=8689688;
 RA Frangiskakis J.M., Exart A.K., Morris C.A., Mervis C.B.,
 RA Bertrand J., Robinson B.F., Klein B.P., Ensing G.J., Everett L.A.,
 RA Green E.D., Proeschel C., Gutowski N.J., Noble M., Atkinson D.L.,
 RA Odenberg S.J., Keating M.T.;
 RA "Lim-kinase1 hemizygosity implicated in impaired visuospatial
 constructive cognition."
 Cell 86:59-69(1996).
 RL Cell 86:59-69(1996).
 CC -1- FUNCTION: MATRIX STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
 CC NUCAL LIAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
 CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
 CC INTO AN EXTENSIBLE 3D NETWORK.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=P15502-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P15502-2; Sequence=VSP_004243;
 CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
 CC -1- DISEASE: Haploinsufficiency of ELN may be the cause of certain
 CC cardiovascular and musculo-skeletal abnormalities observed in
 CC Williams-Beuren syndrome (WBS), a rare developmental disorder. It
 CC is a contiguous gene deletion syndrome involving genes from
 CC chromosome band 7q11.23.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M17282; AAC98394.1; -;
 CC EMBL; M16983; AAC98394.1; JOINED.
 CC EMBL; M17265; AAC98394.1; JOINED.
 CC EMBL; M17266; AAC98394.1; JOINED.
 CC EMBL; M17267; AAC98394.1; JOINED.
 CC EMBL; M17268; AAC98394.1; JOINED.

DR EMBL; M17270; AAC98394.1; JOINED.
 DR EMBL; M17271; AAC98394.1; JOINED.
 DR EMBL; M17272; AAC98394.1; JOINED.
 DR EMBL; M17273; AAC98394.1; JOINED.
 DR EMBL; M17274; AAC98394.1; JOINED.
 DR EMBL; M17275; AAC98394.1; JOINED.
 DR EMBL; M17276; AAC98394.1; JOINED.
 DR EMBL; M17277; AAC98394.1; JOINED.
 DR EMBL; M17278; AAC98394.1; JOINED.
 DR EMBL; M17279; AAC98394.1; JOINED.
 DR EMBL; M17280; AAC98394.1; JOINED.
 DR EMBL; M17281; AAC98394.1; JOINED.
 DR EMBL; M36860; AAA53190.1; -
 DR EMBL; M24782; AAA53190.1; -
 DR EMBL; U62292; AAB17544.1; -
 DR EMBL; X15603; CAA33627.1; -
 DR PIR; A32707; EAHU.
 DR HSSP; P50099; 1ZFU.
 DR Genew; HGNC:3327; ELN.
 DR MIM; 130160; -
 DR MIM; 194050; -
 DR GO; GO:0005578; C:extracellular matrix; TAS.
 DR GO; GO:0005579; C:extracellular space; TAS.
 DR GO; GO:0005601; F:extracellular matrix structural constituent; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0008015; P:circulation; TAS.
 DR GO; GO:0007397; P:hicogenesis and organogenesis; TAS.
 DR GO; GO:0007585; P:respiratory gaseous exchange; TAS.
 DR InterPro; IPR003979; tropoelastin.
 DR PRINTS; PR01500; TROPOLASTIN.
 DR Structural protein; Connective tissue; Repeat; signal;
 KW Williams-Beuren syndrome; Alternative splicing.
 FT SIGNAL 1 26 ELASTIN
 FT CHAIN 27 730
 FT DISULFID 720 725 Missing (in isoform 2).
 FT VARSPLIC 472 477 /FTID=VSP_004243.
 SO SEQUENCE 730 AA; 63260 MW; AB06D15BA567AE46 CRC64;
 Query Match 100.0%; Score 45; DB 1; Length 730;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
 DB 302 AAAAAAAAAA 312

RESULT 5
 ID4_HUMAN STANDARD; PRT; 161 AA.
 AC P47928; Q13005;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE DNA-binding protein inhibitor ID-4.
 GN ID4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN RP
 RP TISSUE=Abdominal adipose tissue;
 RA Kieselring T.L., Christy B.A.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=99087490; Pubmed=9872455;
 RA Rigollet M., Rich T., Gross-Morand M.S., Molina-Gomes D.,
 RA Viegas-Pequignot E., Junien C.;
 RT "cDNA cloning, tissue distribution and chromosomal localization of
 the human ID4 gene.";
 RL DNA Res. 5:309-313(1998).

RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394461; Pubmed=7665172;
 RT Pagliuca A., Bartoli P.C., Saccone S., della Valle G., Lania L.;
 RT "Molecular cloning of ID4, a novel dominant negative helix-loop-helix
 human gene on chromosome 6p21.3-p22.";
 RL Genomics 27:200-203(1995).
 [4]
 RP SEQUENCE FROM N.A.
 RP Mashreghi-Mohammadi M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RX MEDLINE=22388257; Pubmed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshynski S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maris M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC
 CC -1- DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.
 CC -1- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U28368; AAA73923.1; -
 CC EMBL; Y07958; CAA69255.1; -
 CC EMBL; U16153; AAA62882.1; -
 CC EMBL; AL022726; CAA18779.1; -
 CC EMBL; BC014941; AAI14941.1; -
 CC PIR; G01855; G01855.
 DR Genew; HGNC:5363; ID4.
 DR MIM; 600581; -
 DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH_1;
 DR SMART; SM00353; HLH_1;
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS50888; HLH_2; 1.
 KW Nuclear protein.
 FT DOMAIN 39 48
 FT DOMAIN 65 105
 FT DOMAIN 118 124
 FT CONFLICT 10 14
 FT CONFLICT 39 40
 FT CONFLICT 77 79
 POLY-ALA.
 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 POLY-PRO.
 SGRKA -> Q (IN REF. 3).
 AA -> Q (IN REF. 3).
 RTV -> WL (IN REF. 3).

SQ SEQUENCE 161 AA; 16622 MW; 5B14847AE7337339 CRC64;
 Query Match 93.3%; Score 42; DB 1; Length 161;
 Best Local Similarity 90.9%; Pred. No. 6;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAX 11
 |||||
 DB 39 AAAAAAAAAAAR 49

RESULT 6
 ID4_MOUSE
 ID ID4_MOUSE STANDARD; PRT; 161 AA.
 AC P4139;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-binding protein inhibitor ID-4.
 GN ID4 OR ID-4 OR IDB4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C; TISSUE=Bone marrow;
 RX MEDLINE=94188125; PubMed=8139914;
 RA Riechmann V., van Crecchten I., Sablitzky F.;
 RT "The expression pattern of Id4, a novel dominant negative helix-loop-
 RT helix protein, is distinct from Id1, Id2 and Id3."
 RL Nucleic Acids Res. 22:749-755(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95057583; PubMed=9838043;
 RA van Crecchten I., Cinato E., Fox M., King E.R., Newton J.S.,
 RA Riechmann V., Sablitzky F.;
 RT "Structure, chromosomal localization and expression of the murine
 RT dominant negative helix-loop-helix Id4 gene."
 RL Biochim. Biophys. Acta 1443:55-64(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95051333; PubMed=9831657;
 RA Maniani A., Hernandez M.C., Kuo W.-L., Israel M.A.;
 RT "The mouse Id2 and Id4 genes: structural organization and chromosomal
 RT localization."
 RL Gene 222:229-235(1998).
 CC -1- FUNCTION: ID (INHIBITOR OF DNA BINDING) HLH PROTEINS LACK A BASIC
 CC DNA-BINDING DOMAIN BUT ARE ABLE TO FORM HETERODIMERS WITH OTHER
 CC HLH PROTEINS, THEREBY INHIBITING DNA BINDING.
 CC -1- SUBUNIT: HETERODIMER WITH OTHER HLH PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "ID" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X75018; CAAS2926.1; -;
 DR EMBL; AJ001972; CAAS0120.1; -;
 DR EMBL; AF077859; AAD05213.1; -;
 DR PIR; S43260; S43260.
 DR TRANSFAC; T01658; -;
 DR MGD; MGI:99414; Id4.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PFO0010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH_1; 1.

DR PROSITE; PS50888; HLH_2; 1.
 KW Nuclear protein.
 FT DOMAIN 39 48 POLY-ALA.
 FT DOMAIN 65 105 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 118 124 POLY-PRO.
 SQ SEQUENCE 161 AA; 16596 MW; 2DCFA7AF7EE7EED CRC64;

QY 1 AAAAAAAAAAAX 11
 |||||
 DB 39 AAAAAAAAAAAR 49

RESULT 7
 SX21_HUMAN
 ID SX21_HUMAN STANDARD; PRT; 276 AA.
 AC Q9Y651;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcription factor SOX-21.
 GN SOX21.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99373261; PubMed=10441749;
 RA Malas S., Duthe S., Deloukas P., Epiakopou V.;
 RT "The isolation and high-resolution chromosome mapping of human SOX14
 RT and SOX21; two members of the SOX gene family related to SOX1, SOX2,
 RL Mamm. Genome 10:934-937(1999).
 CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR OF TRANSCRIPTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: Contains 1 HMG box domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF107044; AAC95381.1; -;
 DR HSP; Q05066; IHRX.
 DR TRANSFAC; T04921; -;
 DR GeneW; HGNC:11197; SOX21.
 DR MIM; 604974; -;
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
 DR InterPro; IPR000910; HMG_12_box.
 DR Pfam; PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 DR PROSITE; PS50118; HMG_BOX_2; 1.
 KW Transcription regulation; Repressor; DNA-binding; Nuclear protein.
 FT DNA BINDING 8 76 HMG_BOX.
 SQ SEQUENCE 276 AA; 28580 MW; 990C89B7BC9A96B CRC64;

QY 1 AAAAAAAAAAAX 11
 |||||
 DB 137 AAAAAAAAAAAR 147

Query Match 93.3%; Score 42; DB 1; Length 276;
 Best Local Similarity 90.9%; Pred. No. 8.9;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 8
SX21_CHICK STANDARD; PRT; 280 AA.
ID 137 AAAAAAAAAA 11
AC Q62520; STANDARD; PRT; 530 AA.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription factor SOX-21.
GN SOX21.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX MEDLINE=99400103; Pubmed=10473124;
RA Uchikawa M., Kamachi Y., Kondoh H.;
RT "Two distinct subgroups of Group B Sox genes for transcriptional
RT activators and repressors: their expression during embryonic
RT organogenesis of the chicken."
RL Mech. Dev. 84:103-120(1999).
CC -1- FUNCTION: Acts as a negative regulator of transcription.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potentially).
CC -1- TISSUE SPECIFICITY: Expressed predominantly in CNS.
CC -1- SIMILARITY: Contains 1 HMG box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB026623; BAA7266.1; -.
DR HSSP; Q05066; 1HRY.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PSS0118; HMG_BOX_2; 1.
KM Transcription regulation; Repressor; DNA-binding; Nuclear protein.
FT DNA BIND 8 76 HMG_BOX.
FT DOMAIN 137 146 POLY-ALA.
FT DOMAIN 154 166 POLY-ALA.
FT DOMAIN 183 187 POLY-SER.
FT DOMAIN 212 223 POLY-ALA.
SQ SEQUENCE 280 AA; 28796 MW; DA858428B29F1F47 CRC64;

Query Match 93.3%; Score 42; DB 1; Length 280;
Best Local Similarity 90.9%; Pred. No. 9;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
Db 137 AAAAAAAAAA 147

RESULT 9
ZIC2_MOUSE STANDARD; PRT; 530 AA.
ID 024708; STANDARD; PRT; 764 AA.
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase corkscrew (EC 3.1.3.48) (Fragment).
GN ZIC2.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7244;
OX [1]

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=96132643; Pubmed=8557628;
RA Aruga J., Nagai T., Tokuyama T., Hayashizaki Y., Okazaki Y.,
RA Chapman V.M., Mikoshiba K.;
RT "The mouse zic gene family. Homologues of the Drosophila pair-rule
RT gene odd-paired."
RL J. Biol. Chem. 271:1043-1047(1996).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: CNS. A HIGH LEVEL EXPRESSION IS SEEN IN THE
CC CEREBELLUM.
CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 4 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D70848; BAA1115.1; -.
DR HSSP; P08047; ISP2.
DR TRANSFAC; T04670; -.
DR MGD; MGI:106679; Zic2.
DR GO; GO:0007417; P,central nervous system development; IGI.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR Prodom; PD000003; ZnF_C2H2; 1.
DR SMART; SM00355; ZnF_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_2; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_3; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_4; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_5; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_6; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_7; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_8; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_9; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_10; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_11; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_12; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_13; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_14; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_15; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_16; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_17; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_18; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_19; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_20; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_21; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_22; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_23; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_24; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_25; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_26; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_27; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_28; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_29; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_30; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_31; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_32; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_33; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_34; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_35; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_36; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_37; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_38; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_39; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_40; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_41; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_42; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_43; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_44; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_45; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_46; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_47; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_48; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_49; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_50; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_51; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_52; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_53; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_54; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_55; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_56; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_57; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_58; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_59; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_60; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_61; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_62; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_63; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_64; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_65; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_66; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_67; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_68; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_69; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_70; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_71; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_72; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_73; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_74; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_75; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_76; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_77; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_78; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_79; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_80; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_81; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_82; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_83; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_84; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_85; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_86; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_87; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_88; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_89; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_90; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_91; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_92; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_93; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_94; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_95; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_96; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_97; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_98; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_99; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_100; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_101; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_102; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_103; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_104; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_105; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_106; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_107; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_108; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_109; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_110; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_111; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_112; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_113; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_114; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_115; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_116; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_117; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_118; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_119; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_120; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_121; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_122; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_123; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_124; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_125; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_126; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_127; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_128; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_129; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_130; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_131; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_132; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_133; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_134; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_135; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_136; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_137; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_138; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_139; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_140; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_141; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_142; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_143; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_144; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_145; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_146; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_147; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_148; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_149; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_150; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_151; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_152; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_153; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_154; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_155; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_156; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_157; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_158; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_159; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_160; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_161; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_162; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_163; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_164; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_165; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_166; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_167; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_168; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_169; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_170; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_171; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_172; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_173; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_174; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_175; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_176; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_177; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_178; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_179; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_180; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_181; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_182; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_183; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_184; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_185; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_186; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_187; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_188; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_189; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_190; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_191; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_192; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_193; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_194; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_195; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_196; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_197; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_198; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_199; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_200; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_201; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_202; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_203; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_204; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_205; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_206; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_207; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_208; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_209; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_210; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_211; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_212; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_213; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_214; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_215; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_216; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_217; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_218; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_219; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_220; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_221; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_222; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_223; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_224; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_225; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_226; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_227; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_228; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_229; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_230; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_231; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_232; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_233; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_234; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_235; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_236; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_237; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_238; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_239; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_240; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_241; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_242; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_243; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_244; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_245; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_246; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_247; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_248; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_249; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_250; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_251; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_252; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_253; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_254; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_255; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_256; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_257; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_258; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_259; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_260; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_261; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_262; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_263; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_264; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_265; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_266; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_267; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_268; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_269; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_270; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_271; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_272; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_273; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_274; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_275; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_276; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_277; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_278; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_279; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_280; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_281; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_282; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_283; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_284; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_285; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_286; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_287; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_288; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_289; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_290; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_291; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_292; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_293; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_294; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_295; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_296; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_297; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_298; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_299; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_300; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_301; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_302; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_303; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_304; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_305; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_306; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_307; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_308; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_309; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_310; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_311; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_312; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_313; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_314; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_315; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_316; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_317; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_318; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_319; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_320; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_321; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_322; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_323; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_324; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_325; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_326; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_327; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_328; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_329; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_330; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_331; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_332; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_333; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_334; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_335; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_336; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_337; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_338; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_339; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_340; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_341; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_342; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_343; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_344; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_345; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_346; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_347; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_348; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_349; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_350; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_351; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_352; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_353; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_354; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_355; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_356; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_357; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_358; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_359; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_360; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_361; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_362; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_363; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_364; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_365; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_366; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_367; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_368; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_369; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_370; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_371; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_372; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_373; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_374; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_375; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_376; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_377; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_378; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_379; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_380; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_381; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_382; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_383; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_384; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_385; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_386; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_387; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_388; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_389; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_390; 4.
DR PROSITE; PSS0157; ZINC_FINGER_C2H2_391; 4.

```

RP SEQUENCE FROM N.A.
 RA Melnick M.B., Melnick C.B., Larsen I., Perrimon N., Perkins I.A.;
 RT "The role of the Drosophila corkscrew protein as a transducer
 RL downstream of receptor tyrosine kinases is functionally conserved.";
 CC Submitted (MAR-1995) to the EMBL/Genbank/DBJ database.
 CC -1- FUNCTION: Required in all receptor tyrosine kinase signaling
 CC pathways. Functions downstream of the receptor tyrosine kinase
 CC torso, acting in concert with D-Raf via tailless. Also functions
 CC downstream of Egfr (epidermal growth factor receptor) and btl
 CC (fibroblast growth factor receptor). The SH2 domain suggests that
 CC csw effects its role by mediating heteromeric protein
 CC interactions. Maternally required for normal determination of cell
 CC fates at the termini of the embryo. Required for cell fate
 CC specification of the ventral ectoderm, in the developing embryonic
 CC CNS and for embryonic tracheal cell migration. Functions during
 CC imaginal development for proper formation of adult structures such
 CC as eyes, arista, L5 wing vein and the tarsal claw (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLASS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: Contains at least 1 SH2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U22356; AAB02545.1; --
 CC HSSP: Q06124; ZSPH.
 CC DR FlyBase: FBgn0016496; Dvir\csw.
 CC DR InterPro: IPR000980; SH2.
 CC DR InterPro: IPR000387; TYR_PTPase.
 CC DR InterPro: IPR000242; TYR_PP.
 CC DR Pfam: PF00017; SH2; 1.
 CC DR Pfam: PF00102; Y_phosphatase; 1.
 CC DR PRINTS: PR00700; PRTYPHPTASE.
 CC DR PRINTS: PR00401; SH2DOMAIN.
 CC DR ProDom: PD000093; SH2; 1.
 CC DR SMART: SM00194; PTPc; 1.
 CC DR SMART: SM00252; SH2; 1.
 CC DR PROSITE: PS50001; SH2; 1.
 CC DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 CC DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 CC DR PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.
 CC DR Hydrolyase; SH2 domain; Developmental protein.
 CC FT NON_TER 1 1
 CC FT DOMAIN 1 95 SH2.
 CC FT DOMAIN 117 522 PROTEIN-TYROSINE PHOSPHATASE.
 CC FT DOMAIN 174 325 PTPASE INSERT (CYS/SER-RICH).
 CC FT ACT_SITE 460 460 BY SIMILARITY.
 CC FT DOMAIN 186 191 POLY-SER.
 CC FT DOMAIN 278 281 POLY-ALA.
 CC FT DOMAIN 576 600 ALA-RICH.
 CC FT DOMAIN 613 616 POLY-ASN.
 CC FT DOMAIN 617 656 SER-RICH.
 CC FT DOMAIN 687 753 POLY-GLN.
 CC FT DOMAIN 694 753 ALA-RICH.
 CC FT DOMAIN 755 761 POLY-PRO.
 CC SQ SEQUENCE 764 AA; 82125 MW; 66008EAB560A2F7D CRC64;
 QY Query Match 93.3%; Score 42; DB 1; Length 764;
 Db Best Local Similarity 90.9%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
 HUES_DROME STANDARD; PRT; 1077 AA.
 AC Q02308; Q9VDK0; Q9VDK1;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hairless protein.
 GN H OR CG5460.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephyrididae; Drosophilidae; Drosophila.
 OK NCBI_Taxid=7227;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=92387549; PubMed=1516831;
 RA Bang A.G., Posakony J.W.;
 RT "The Drosophila gene Hairless encodes a novel basic protein that
 RL controls alternative cell fates in adult sensory organ development.";
 RN 12
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=93041287; PubMed=1419850;
 RA Maier D., Stumm G., Kuhn K., Preiss A.;
 RT "Hairless, a Drosophila gene involved in neural development, encodes
 RL a novel, heparin rich protein.";
 RN 3
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazek R.G., Chang M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrita J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mayhew A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpem G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Modarres C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson K.A.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheelers F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitak R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RC Science 287:2185-2195 (2000).
 RL -1- FUNCTION: IS A POTENT ANTAGONIST OF NEUROGENIC GENE ACTIVITY

CC DURING SENSORY ORGAN DEVELOPMENT. THE EXPRESSION OF DISTINCT CELL
 CC FATES BY THE TRICHOGEN (SHAFT) / TOMOGEN (SOCKET) SISTER CELL
 CC PAIR DEPENDS ON THE LEVEL OF H ACTIVITY. A CERTAIN THRESHOLD LEVEL
 CC OF H ACTIVITY IS REQUIRED, BELOW WHICH BOTH SISTER CELLS ADOPT THE
 CC TOMOGEN FATE.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q02308-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q02308-2; Sequence=VSP_006952;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: OVARY, EMBRYOS, LARVAL AND PUPAL IMAGINAL
 CC DISCS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M95192; AAA28607.1; ALT INIT.
 CC DR EMBL; X67239; CAA47664.1; -.
 CC DR EMBL; AE003731; AAF55790.1; -.
 CC DR EMBL; AB003731; AAF55791.1; -.
 CC DR PIR; A44067; A44067.
 CC DR FlyBase; FBgn0001169; H.
 CC DR GO; GO:0003714; P:Transcription co-repressor activity; IDA.
 CC DR GO; GO:0007219; P:N receptor signaling pathway; NAS.
 CC DR GO; GO:0008052; P:sensory organ determination; IMP.
 CC KW Developmental protein; Nuclear protein; DNA-binding;
 CC Alternative splicing.
 CC FT DOMAIN 115 123 THR-RICH.
 CC FT DOMAIN 642 648 POLY-SER.
 CC FT DOMAIN 879 891 POLY-ALA.
 CC FT DOMAIN 937 946 POLY-ALA.
 CC FT DOMAIN 964 974 ALA-RICH.
 CC FT DOMAIN 979 1008 HIS/PRO-RICH (PRO MOTIF).
 CC FT VARSPLIC 1 18 Missing (in isoform 2).
 CC FT CONFLICT 151 151 /FtId=VSP_006952.
 CC FT CONFLICT 680 680 S -> A (IN REF. 2).
 CC FT CONFLICT 702 703 S -> F (IN REF. 3).
 CC FT CONFLICT 891 891 QH -> LL (IN REF. 2).
 CC FT CONFLICT 964 967 A -> R (IN REF. 2).
 CC FT CONFLICT 974 974 AAVA -> RLUP (IN REF. 2).
 CC FT CONFLICT 974 974 MISSING (IN REF. 2).
 CC SQ SEQUENCE 1077 AA; 111039 MW; A94BF1A27579E2F1 CRC64;
 CC -----
 CC Query Match 93.3%; Score 42; DB 1; Length 1077;
 CC Best Local Similarity 90.9%; Pred. No. 25;
 CC Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC Oy 1 AAAAAAAAAAAX 11
 CC Db 937 AAAAAAAAAAAR 947

OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92249205; PubMed=1576962;
 RA Macdonald P.M.;
 RT "The Drosophila pumilio gene: an unusually long transcription unit
 RT and an unusual protein.";
 RL Development 114:221-234(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93093466; PubMed=1459455;
 RA Barker D.D., Wang C., Moore J., Dickinson L.K., Lehmann R.;
 RT "Pumilio is essential for function but not for distribution of the
 RT Drosophila abdominal determinant Nanos.";
 RL Genes Dev. 6:2312-2326(1992).
 CC -1- FUNCTION: PUM IS THE ONLY GENE REQUIRED FOR NOS ACTIVITY THAT IS
 CC NOT ALSO REQUIRED FOR POSTERIOR LOCALIZATION OF GERM LINE
 CC DETERMINANTS. PUM IS REQUIRED DURING EMBRYOGENESIS WHEN NOS
 CC ACTIVITY APPARENTLY MOVES ANTERIORLY FROM THE POSTERIOR POLE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. IT IS CONCENTRATED IN THE
 CC CORTICAL REGION OF THE EMBRYO BENEATH THE NUCLEI.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE OVARIES AND DURING THE
 CC EMBRYOGENESIS.
 CC -1- DOMAIN: CONSISTS MAINLY OF REGIONS ENRICHED IN A SINGLE AMINO
 CC ACID.
 CC -1- DISEASE: LETHAL DEFECTIVE IN POSTERIOR PATTERN FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE PUMILIO/MP75 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X62589; CAA44474.1; -.
 CC DR EMBL; L07943; AAB59189.1; -.
 CC DR PIR; A46221; A46221.
 CC DR FlyBase; FBgn0003165; pum.
 CC DR GO; GO:0008258; P:head involution; IMP.
 CC DR InterPro; IPR001313; Pumilio/Put.
 CC DR SMART; SM00025; Pumilio; 8.
 CC KW Developmental protein; Repeat.
 CC FT DOMAIN 34 45 ALA-RICH.
 CC FT DOMAIN 57 77 GLY/VNL-RICH.
 CC FT DOMAIN 83 93 ALA-RICH.
 CC FT DOMAIN 130 174 GLY-RICH.
 CC FT DOMAIN 152 164 POLY-GLY.
 CC FT DOMAIN 181 212 POLY-GLY.
 CC FT DOMAIN 213 236 GLN-RICH.
 CC FT DOMAIN 262 274 POLY-GLN.
 CC FT DOMAIN 571 599 GLY-RICH.
 CC FT DOMAIN 708 725 GLY-RICH.
 CC FT DOMAIN 936 946 POLY-GLN.
 CC FT DOMAIN 1050 1062 POLY-ALA.
 CC FT DOMAIN 1111 1326 6 X 36 AA APPROXIMATE TANDEM REPEATS.
 CC FT REPEAT 1111 1146 1.
 CC FT REPEAT 1147 1182 2.
 CC FT REPEAT 1183 1218 3.
 CC FT REPEAT 1219 1254 4.
 CC FT REPEAT 1255 1290 5.
 CC FT REPEAT 1291 1326 6.
 CC FT CONFLICT 362 362 A -> S (IN REF. 2).
 CC FT CONFLICT 1103 1103 R -> P (IN REF. 2).
 CC FT CONFLICT 1406 1407 KN -> PH (IN REF. 2).
 CC FT CONFLICT 1496 1496 V -> I (IN REF. 2).
 CC FT CONFLICT 1519 1519 S -> G (IN REF. 2).
 CC SQ SEQUENCE 1533 AA; 157521 MW; C453A2321B8BDBDC CRC64;
 CC -----
 CC Query Match 93.3%; Score 42; DB 1; Length 1533;
 CC Best Local Similarity 90.9%; Pred. No. 32;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 937 AAAAAAAAAAR 947

RESULT 13
 ANP LIMFE STANDARD; PRT; 97 AA.

AC P09031;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE Antifreeze protein precursor (AFP).
 OS Limanda ferruginea (Yellowtail flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectidae; Pleuronectidae; Limanda.
 NCBI_TaxId=8256;
 RX MEDLINE=88029483; PubMed=3665937;
 RA Scott G.K., Davies P.L., Shears M.A., Fletcher G.L.;
 RT "Structural variations in the alanine-rich antifreeze proteins of the
 RL pleuronectinae";
 RT Eur. J. Biochem. 168:629-633(1987).
 CC -1- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-1 AFP FAMILY. TYPE 1 AFP ARE
 CC ALANINE-RICH, AMPHIPHILIC AND ALPHA-HELICAL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X06356; CAA29655.1; -
 DR PIR; S02376; S02376.
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PR00308; ANTIFREEZE1.
 KM Antifreeze protein; Repeat; Signal.
 FT SIGNAL 1 23
 FT PROPEP 24 48
 FT CHAIN 49 97
 FT SEQUENCE 97 AA; 8865 MW; 62AD582DF8E459B6 CRC64;
 (PROBABLE).
 ANTI-FREEZE PROTEIN.
 Query Match 91.1%; Score 41; DB 1; Length 97;
 Best Local Similarity 90.9%; Pred. No. 5.6;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 57 AAAAAAAAAAK 67

RESULT 14
 ASH1 MOUSE STANDARD; PRT; 231 AA.

AC Q02067;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Achaete-scute homolog 1 (Mash-1).
 GN ASCL1 OR ASH1 OR MASH1 OR MASH-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;

RM [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93144349; PubMed=8424959;
 RA del Amo F., Gendron-Magnire M., Gridley T.;
 RT "Cloning, sequencing and expression of the mouse mammalian
 RT achaete-scute homolog 1 (MASH1).";
 RL Biochim. Biophys. Acta 1171:323-327(1993).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94031847; PubMed=8217843;
 RA Guillemot F., Joyner A.L.;
 RT "Dynamic expression of the murine Achaete-Scute homologue Mash-1 in
 RT the developing nervous system.";
 RL Mech. Dev. 42:171-185(1993).
 CC -1- FUNCTION: MAY PLAY A ROLE AT EARLY STAGES OF DEVELOPMENT OF
 CC SPECIFIC NEURAL LINEAGES IN MOST REGIONS OF THE CNS, AND OF
 CC SEVERAL LINEAGES IN THE PNS. ESSENTIAL FOR THE GENERATION OF
 CC OLFACTORY AND AUTONOMIC NEURONS. ACTIVATES TRANSCRIPTION BY
 CC BINDING TO THE E BOX (5'-CANNNG-3').
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. FORMS A HETERODIMER WITH E12/E47.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: DEVELOPING CNS AND PNS AT EMBRYONIC AND POST-
 CC NATAL STAGES.
 CC -1- DEVELOPMENTAL STAGE: BETWEEN EMBRYONIC DAYS 8.5 AND 10.5 IT IS
 CC FOUND IN THE NEUROEPITHELIUM OF THE MIDBRAIN AND VENTRAL
 CC FOREBRAIN, AS WELL AS IN THE SPINAL CORD. BETWEEN DAYS 10.5 AND
 CC 12.5 ITS EXPRESSION PATTERN CHANGES FROM A RESTRICTED TO A
 CC WIDESPREAD ONE, IT IS THEN FOUND AT VARIABLE LEVELS IN THE
 CC VENTRICULAR ZONE IN ALL REGIONS OF THE BRAIN. FROM DAY 12.5 TO
 CC POST-NATAL STAGES IT IS ALSO EXPRESSED IN CELLS OUTSIDE OF THE
 CC VENTRICULAR ZONE THROUGH THE BRAIN, AND IN ADDITION IT IS ALSO
 CC EXPRESSED DURING DEVELOPMENT OF THE OLFACTORY EPITHELIUM AND
 CC NEURAL RETINA.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. ASC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M95603; AAA37780.1; -
 DR PIR; S28186; S28186.
 DR TRANSFAC; T01619; -
 DR MGD; MGI:96919; Ascl1.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:007400; P:neuroblast cell fate determination; IMP.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH_1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00038; HLH_1; 1.
 DR PROSITE; PS50888; HLH_2; 1.
 KM Neurogenesis; Differentiation; Developmental protein; DNA-binding;
 KW Nuclear protein.
 FT DOMAIN 30 43
 FT DOMAIN 47 52
 FT DNA BIND 116 126
 FT DOMAIN 127 166
 FT SEQUENCE 231 AA; 24755 MW; 5FAB4A62052ABF7 CRC64;
 (POLY-ALA.
 BASIC DOMAIN.
 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 Query Match 91.1%; Score 41; DB 1; Length 231;
 Best Local Similarity 90.9%; Pred. No. 11;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAK 11
 |||||
 Db 34 AAAAAAAAAAO 44

Search completed: February 19, 2004, 10:03:42
Job time : 12 secs

```
RESULT 15
ASH1_RAT STANDARD; PRT; 233 AA.
ID ASH1_RAT
AC p19359;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Achaeete-scute homolog 1.
GN ASCL1 OR ASH1 OR MASH-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=90363294; PubMed=2392153;
RA Johnson J.E., Birren S.J., Anderson D.J.;
RT "Two rat homologues of Drosophila achaeete-scute specifically
RL Nature 346:858-861(1990).
CC -1- FUNCTION: MAY PLAY A ROLE AT EARLY STAGES OF DEVELOPMENT OF
CC SPECIFIC NEURAL LINEAGES IN MOST REGIONS OF THE CNS, AND OF
CC SEVERAL LINEAGES IN THE PNS. ESSENTIAL FOR THE GENERATION OF
CC OLFACTORY AND AUTONOMIC NEURONS. ACTIVATES TRANSCRIPTION BY
CC BINDING TO THE E BOX (5'-CANNYG-3').
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. FORMS A HETERODIMER WITH E12/E47.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: DEVELOPING CNS AND PNS AT EMBRYONIC AND POST-
CC NATAL STAGES.
CC -1- DEVELOPMENTAL STAGE: IT IS FIRST OBSERVED AFTER NEURULATION, IN
CC 10.5 DAY OLD (E10.5) RAT EMBRYOS, AND IS RESTRICTED TO SUBSETS OF
CC NEUROEPITHELIAL CELLS IN THE SPINAL CORD AND THE BRAIN. BETWEEN
CC E10.5 AND E13.5. IN THE PERIPHERY, ITS EXPRESSION IS RESTRICTED TO
CC SOME LINEAGES OF NEURAL CREST-DERIVED CELLS, NAMELY IN SYMPATHETIC
CC AND ENTERIC NEURAL PRECURSORS. IN THE PNS ITS EXPRESSION IS
CC EXTINGUISHED AT OR BEFORE DIFFERENTIATION.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS. ASG SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X53725; CAA37760.1; -.
CC DR PIR; S11563; S11563.
CC DR TRANSFAC; T00484; -.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH_1.
CC DR SMART; SM00353; HLH_1.
CC DR PROSITE; PS00038; HLH_1; 1.
CC DR PROSITE; PS50888; HLH_2; 1.
CC KW Neurogenesis; Differentiation; Developmental protein; DNA-binding;
CC Nuclear protein.
CC FT DOMAIN 30 44 POLY-ALA.
CC FT DOMAIN 48 54 POLY-GLN.
CC FT DNA BIND 118 128 BASIC DOMAIN.
CC FT DOMAIN 129 168 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
CC SQ SEQUENCE 233 AA; 24972 MW; 036BDAC8E2D3274 CRC64;
Query Match 91.1%; Score 41; DB 1; Length 233;
Best Local Similarity 90.9%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AAAAAAAAAAK 11
DB 35 AAAAAAAAAAQ 45

QY	1	AAAAAAAAAAK	11
CH2			
DB	13	AAAAAAAAAAK	23
RESULT 2			
Q8CH22			
ID	Q8CH22	PRELIMINARY;	PRT;
AC	Q8CH22;		171 AA

DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Basic protein CTR1.
 GN CTR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; Tissue=Testis; Escalier D., McFarland L., Xu P.-X.;
 RA Xu X., Bai X., Silvius D., Escalier D., McFarland L., Xu P.-X.;
 RT "CK2 differentially phosphorylate a family of novel spermatid-specific
 basic nuclear proteins."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF463500; AAO15673.1; -
 SQ SEQUENCE 171 AA; 19175 MW; 24829F50B121B68 CRC64;
 Query Match 100.0%; Score 45; DB 11; Length 171;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAA 11
 Db 13 AAAAAAAAAA 23

RESULT 3
 Q8S4Y0 PRELIMINARY; PRT; 396 AA.
 ID O8S4Y0;
 AC O8S4Y0;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Putative ERBBP-type transcription factor.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 NX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Park M., Moon E., Hwang D.-J.;
 RT "Molecular cloning of a putative ERBBP-type transcription factor in
 rice."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF364176; AAM00285.1; -
 DR Gramene; Q8S4Y0; -
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR ProDom; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 396 AA; 42597 MW; EC3FD53BFF5F9361 CRC64;
 Query Match 100.0%; Score 45; DB 10; Length 396;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAA 11
 Db 15 AAAAAAAAAA 25

RESULT 4
 Q8N1Z3 PRELIMINARY; PRT; 397 AA.
 ID O8N1Z3;
 AC O8N1Z3;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ37192.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Alzheimer cortex;
 RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawal-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuna M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshina A., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuno Y., Nagai K., Isoyagi T.;
 RT "NEDD human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 IONS.
 DR EMBL, AK094511; BAC04371.1; -
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR001781; LIM.
 DR InterPro; IPR007107; LIM_homeo.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00412; LIM; 2.
 DR ProDom; PD000010; Homeobox; 1.
 DR ProDom; PD000094; LIM; 2.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00132; LIM; 2.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 DR PROSITE; PS00478; LIM_DOMAIN 1; 2.
 DR PROSITE; PS50023; LIM_DOMAIN 2; 2.
 DR PROSITE; PS50558; LIM_HOMEDOMAIN; 1.
 KM Hypothetical protein; LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 397 AA; 43363 MW; D4DB2956EA6478D CRC64;
 Query Match 100.0%; Score 45; DB 4; Length 397;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAA 11
 Db 178 AAAAAAAAAA 188

RESULT 5
 Q8GSB1 PRELIMINARY; PRT; 594 AA.
 ID Q8GSB1;
 AC Q8GSB1;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Putative calcium/calmodulin-dependent protein kinase CamK.
 GN P0524G08.10 OR OJ1340 C08.39.
 OS Oryza sativa (Japanese cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 NX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
 clone:P0524G08."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
 clone:OJ1340 C08."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AP004671; BAC16472.1; -
 DR EMBL, AP005292; BAC45222.1; -

```

KW Kinase.
SQ SEQUENCE 594 AA; 65242 MW; 46FF4E13CEFD7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 45; DB 10; Length 594;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
   |||||
   |||||
Db 36 AAAAAAAAAAAK 46

RESULT 6
ID 015337 PRELIMINARY; PRT; 602 AA.
AC 015337;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Elastin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Li D.Y., Toland A.E., Boak B.B., Atkinson D.L., Ensing G.J.,
RT "Elastin point mutations cause an obstructive vascular disease,
supravalvular aortic stenosis."
Hum. Mol. Genet. 0:0-0(1997).
DR EMBL; U93037; AAB65620.1; JOINED.
DR EMBL; U93034; AAB65620.1; JOINED.
DR EMBL; U93036; AAB65620.1; JOINED.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR003979; tropoelastin.
DR PRINTS; PR01500; TROPOELASTIN.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
FT NON_TER 1
FT NON_TER 602
SQ SEQUENCE 602 AA; 51807 MW; 53B5B9A71EF04807 CRC64;

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 602;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
   |||||
   |||||
Db 274 AAAAAAAAAAAK 284

RESULT 7
ID 015336 PRELIMINARY; PRT; 635 AA.
AC 015336;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Elastin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Li D.Y., Toland A.E., Boak B.B., Atkinson D.L., Ensing G.J.,
RT "Elastin point mutations cause an obstructive vascular disease,
supravalvular aortic stenosis."
Hum. Mol. Genet. 0:0-0(1997).
DR EMBL; U93037; AAB65621.1; JOINED.
DR EMBL; U93034; AAB65621.1; JOINED.

```

```

DR EMBL; U93035; AAB65621.1; JOINED.
DR EMBL; U93036; AAB65621.1; JOINED.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR003979; tropoelastin.
DR PRINTS; PR01500; TROPOELASTIN.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
FT NON_TER 1
FT NON_TER 635
SQ SEQUENCE 635 AA; 55279 MW; 72950C364127B2A4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 635;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
   |||||
   |||||
Db 274 AAAAAAAAAAAK 284

RESULT 8
ID 08NB14 PRELIMINARY; PRT; 643 AA.
AC 08NB14;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein NT2RP3003474.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Oca T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA Nagahari K., Sugano S., Isogai T.,
RT "HRI human cDNA sequencing project."
Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK075554; U011696.1; -.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR000158; Flaz.
DR InterPro; IPR003979; tropoelastin.
DR PRINTS; PR00423; CELDVISFTSZ.
DR PRINTS; PR01500; TROPOELASTIN.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
KM Hypothetical protein.
SQ SEQUENCE 643 AA; 55629 MW; FDFC042617E72A69 CRC64;

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 643;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
   |||||
   |||||
Db 266 AAAAAAAAAAAK 276

RESULT 9
ID 014235 PRELIMINARY; PRT; 687 AA.
AC 014235;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Elastin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=87274906; PubMed=3038460;

```

RA Indik Z., Yoon K., Morrow S.D., Cicilia G., Rosenbloom J.,
 RA Rosenbloom J., Ornstein-Goldstein N.;
 RT "Structure of the 3' region of the human elastin gene: great abundance
 RT of Alu repetitive sequences and few coding sequences.";
 RN Connect. Tissue Res. 16:197-211(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87289668; PubMed=3039501;
 RA Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Peltonen L., Rosenbloom J.;
 RT "Alternative splicing of human elastin mRNA indicated by sequence
 RT analysis of cloned genomic and complementary DNA.";
 RN Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
 DR EMBL; M17282; AAC98393.1; JOINED.
 DR EMBL; M16983; AAC98393.1; JOINED.
 DR EMBL; M17265; AAC98393.1; JOINED.
 DR EMBL; M17266; AAC98393.1; JOINED.
 DR EMBL; M17267; AAC98393.1; JOINED.
 DR EMBL; M17268; AAC98393.1; JOINED.
 DR EMBL; M17271; AAC98393.1; JOINED.
 DR EMBL; M17272; AAC98393.1; JOINED.
 DR EMBL; M17273; AAC98393.1; JOINED.
 DR EMBL; M17275; AAC98393.1; JOINED.
 DR EMBL; M17276; AAC98393.1; JOINED.
 DR EMBL; M17277; AAC98393.1; JOINED.
 DR EMBL; M17278; AAC98393.1; JOINED.
 DR EMBL; M17279; AAC98393.1; JOINED.
 DR EMBL; M17281; AAC98393.1; JOINED.
 DR HSSP; P50099; 1ZFU.
 DR InterPro; IPR001179; FKBP_PPIase.
 DR InterPro; IPR003979; tropoelastin.
 DR PRINTS; PR01500; TROPOELASTIN.
 DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
 SQ SEQUENCE 687 AA; 59529 MW; 864068C4C8B9F88F CRC64;

Query Match 100.0%; Score 45; DB 4; Length 687;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
 DB 302 AAAAAAAAAA 312

RESULT 10
 Q14234 PRELIMINARY; PRT; 757 AA.
 ID Q14234
 AC Q14234;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Elastin.
 GN ELN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87274906; PubMed=3038460;
 RA Indik Z., Yoon K., Morrow S.D., Cicilia G., Rosenbloom J.,
 RA Rosenbloom J., Ornstein-Goldstein N.;
 RT "Structure of the 3' region of the human elastin gene: great abundance
 RT of Alu repetitive sequences and few coding sequences.";
 RN Connect. Tissue Res. 16:197-211(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87289668; PubMed=3039501;
 RA Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Peltonen L., Rosenbloom J.;
 RT "Alternative splicing of human elastin mRNA indicated by sequence
 RT analysis of cloned genomic and complementary DNA.";
 RN Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).

DR EMBL; M17282; AAC98395.1; JOINED.
 DR EMBL; M16983; AAC98395.1; JOINED.
 DR EMBL; M17265; AAC98395.1; JOINED.
 DR EMBL; M17266; AAC98395.1; JOINED.
 DR EMBL; M17267; AAC98395.1; JOINED.
 DR EMBL; M17268; AAC98395.1; JOINED.
 DR EMBL; M17270; AAC98395.1; JOINED.
 DR EMBL; M17271; AAC98395.1; JOINED.
 DR EMBL; M17272; AAC98395.1; JOINED.
 DR EMBL; M17273; AAC98395.1; JOINED.
 DR EMBL; M17274; AAC98395.1; JOINED.
 DR EMBL; M17275; AAC98395.1; JOINED.
 DR EMBL; M17276; AAC98395.1; JOINED.
 DR EMBL; M17277; AAC98395.1; JOINED.
 DR EMBL; M17278; AAC98395.1; JOINED.
 DR EMBL; M17279; AAC98395.1; JOINED.
 DR EMBL; M17280; AAC98395.1; JOINED.
 DR EMBL; M17281; AAC98395.1; JOINED.
 DR InterPro; IPR001179; FKBP_PPIase.
 DR InterPro; IPR003979; tropoelastin.
 DR PRINTS; PR01500; TROPOELASTIN.
 DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
 SQ SEQUENCE 757 AA; 66136 MW; 23B7FE5B8AF85CA8 CRC64;

Query Match 100.0%; Score 45; DB 4; Length 757;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAA 11
 DB 302 AAAAAAAAAA 312

RESULT 11
 Q9W2V2 PRELIMINARY; PRT; 804 AA.
 ID Q9W2V2
 AC Q9W2V2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG32683 protein.
 GN CG32683 OR CG2881 OR CG2883 OR CG15303.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter R.G., Helt G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Bocham M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslcek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegum C.,
 RA Jalali M., Kalish W., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodita C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Putative Rhs-related protein.
 GN RSP0151 OR RS02969.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Plasmid megaplasmid.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Ralstoniaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=2161879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Aizat M., Billault A., Brotter P., Camus J.C., Catolico L.,
 RA Chandler M., Choiane N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lave M., Moisan A., Robert C., Sarrin W., Schiex T.,
 RA Siguler P., Thebaud P., Whalen M., Winkler P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646076; CAD17302.1; -;
 DR InterPro: IPR005074; Peptidase_C39.
 DR InterPro: IPR006530; YD.
 DR Pfam: PF03412; Peptidase_C39; 1.
 DR TIGRPFAMs: TIGR01643; YD repeat_2x; 14.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 1710 AA; 187661 MW; AACDBE9646AAE29 CRC64;

Query Match 100.0%; Score 45; DB 16; Length 1710;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
 DB 1605 AAAAAAAAAAAK 1615

RESULT 14
 O8CH17 PRELIMINARY; PRT; 161 AA.
 AC O8CH17;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Helix-loop-helix protein.
 GN ID4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RA Shan L., Yu M., Snyderwine E.G.;
 RT "Cloning and functional analysis of rat Id4 gene."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF468681; AA015695.1; -;
 SQ SEQUENCE 161 AA; 16621 MW; 2AB9E2D69C3909ED CRC64;

Query Match 93.3%; Score 42; DB 11; Length 161;
 Best Local Similarity 90.9%; Pred. No. 44;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
 DB 39 AAAAAAAAAAAK 49

RESULT 15
 O8VH35 PRELIMINARY; PRT; 276 AA.
 AC O8VH35;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)

DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DE 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 GN HMG-box protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/cJ;
 RA Uvanogho D.A.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069926; AL49967.1; -;
 DR InterPro: IPR000104; Antifreeze_1.
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR PRINTS: PR00308; ANTI-FREEZE1.
 DR SMART: SM00398; HMG; 1.
 SQ SEQUENCE 276 AA; 28593 MW; 24A31E0FE24A9078 CRC64;

Query Match 93.3%; Score 42; DB 11; Length 276;
 Best Local Similarity 90.9%; Pred. No. 70;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAAAK 11
 DB 137 AAAAAAAAAAAK 147

Search completed: February 19, 2004, 10:04:30
 Job time : 37 secs